PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C07H 21/02, 21/04, C12Q 1/68 — A1

(11) International Publication Number: WO 96/29338

(43) International Publication Date: 26 September 1996 (26.09.96)

(21) International Application Number: PCT/US96/03896

(22) International Filing Date: 22 March 1996 (22.03.96)

(30) Priority Data:
08/410,540 23 March 1995 (23.03.95) US

(60) Parent Application or Grant
(63) Related by Continuation
US 08/410,540 (CIP)
Filed on 23 March 1995 (23.03.95)

(71) Applicants *(for all designated States except US)*: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA [US/US]; 22nd floor, 300 Lakeside Drive, Oakland, CA 94612-3550 (US). THE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA [US/US]; Suite 300, 3700 Market Street, Philadelphia, PA 19104-3147 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: MILLER, Walter, L. [US/US]; 1683 8th Avenue, San Francisco, CA 94122 (US). LIN, Dong [CN/US]; 1672 10th Avenue, San Francisco, CA 94122 (US). STRAUSS, Jerome, F., III [US/US]; 805 E. Giravers Lane, Wyndmoor, PA 19038 (US).

(74) Agent: BERLINER, Robert; Robbins, Berliner & Carson, 5th floor, 201 N. Figueroa Street, Los Angeles, CA 90012-2628 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: IDENTIFICATION OF GENE MUTATIONS ASSOCIATED WITH CONGENITAL LIPOID ADRENAL HYPERPLASIA

(57) Abstract

An isolated DNA or RNA molecule, wherein said molecule contains: (1) a first sequence consisting of hStAR cDNA, hStAR genomic DNA, or hStAR pseudogene DNA as set forth in Figure 1, Table 6, or Table 7; (2) a second sequence, wherein said second sequence is a subsequence of said first sequence at least 10 nucleotides in length; (3) a third sequence in which at least one nucleotide of said first or second sequence is replaced by a different nucleotide; (4) a fourth sequence in which at least one nucleotide is deleted from or inserted into said first or second sequence; or (5) a fifth sequence complementary to any of said first second, or third sequences; with the provisos that (1) said molecule can be an RNA molecule in which U replaces T in any of said sequences (1) - (5), (2) said third sequence is at least 95 % identical to said first or second sequence, (3) said second sequence is not present in mouse StAR cDNA, and (4) said fourth sequence contains no more than 20 inserted nucleotides and no more than 200 deleted nucleotides.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	**GB**	United Kingdom	**MW**	Malawi
AT	Austria	**GE**	Georgia	**MX**	Mexico
AU	Australia	**GN**	Guinea	**NE**	Niger
BB	Barbados	**GR**	Greece	**NL**	Netherlands
BE	Belgium	**HU**	Hungary	**NO**	Norway
BF	Burkina Faso	**IE**	Ireland	**NZ**	New Zealand
BG	Bulgaria	**IT**	Italy	**PL**	Poland
BJ	Benin	**JP**	Japan	**PT**	Portugal
BR	Brazil	**KE**	Kenya	**RO**	Romania
BY	Belarus	**KG**	Kyrgystan	**RU**	Russian Federation
CA	Canada	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CF	Central African Republic			**SE**	Sweden
CG	Congo	**KR**	Republic of Korea	**SG**	Singapore
CH	Switzerland	**KZ**	Kazakhstan	**SI**	Slovenia
CI	Côte d'Ivoire	**LI**	Liechtenstein	**SK**	Slovakia
CM	Cameroon	**LK**	Sri Lanka	**SN**	Senegal
CN	China	**LR**	Liberia	**SZ**	Swaziland
CS	Czechoslovakia	**LT**	Lithuania	**TD**	Chad
CZ	Czech Republic	**LU**	Luxembourg	**TG**	Togo
DE	Germany	**LV**	Latvia	**TJ**	Tajikistan
DK	Denmark	**MC**	Monaco	**TT**	Trinidad and Tobago
EE	Estonia	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**UG**	Uganda
FI	Finland	**ML**	Mali	**US**	United States of America
FR	France	**MN**	Mongolia	**UZ**	Uzbekistan
GA	Gabon	**MR**	Mauritania	**VN**	Viet Nam

IDENTIFICATION OF GENE MUTATIONS ASSOCIATED WITH CONGENITAL LIPOID ADRENAL HYPERPLASIA

This application is a continuation-in-part of Patent Application Serial Number 08/410,540, which is herein incorporated by reference.

ACKNOWLEDGEMENTS

This invention was supported in part by NIH Grants HD 06274, HD 07688, HD 17481, RR000847, DK 37922, and DK 42154. The U.S. Government has rights in this invention.

INTRODUCTION

Technical Field

This invention is directed to a genetic sequence that has been identified as the locus of mutations that cause congenital lipoid adrenal hyperplasia (lipoid CAH) and to methods for the diagnosis of this disease and for the detection of the presence of the mutated gene as an indication of potential for genetic transmission of the disease.

Background

Steroid hormone synthesis is greatly increased in response to tropic hormone stimulation. Although increased transcription of genes encoding steroidogenic enzymes is important in the chronic hormonal response, the rate-limiting step in the acute response is the transport of cholesterol into mitochondria (J.F. Crivello et al., *J. Biol. Chem,* **255**, 8144 (1980); C.R. Jefcoate et al., *J. Steroid Biochem.* **27**, 721 (1987)). Several molecules have been proposed to participate in this transport, but their roles have not been definitively established.

Early studies showed that congenital lipoid adrenal hyperplasia (lipoid CAH) was an autosomal recessive disorder (Prader, et al. *Helv. Paed Acta* 17,

285-289 (1962)) characterized by a severe deficiency of adrenal and gonadal steroid hormones (H.J. Degenhart et al., *Acta Endocrinol.* **71**, 215 (1972); S. Koizumi *et al., Clin. Chem. Acta.* **77**, 301 (1977); B.P. Hauffa et al., *Clin. Endocrinol.* **23**, 481 (1985)). Affected infants die from salt loss, hyperkalemic acidosis and dehydration unless treated with steroid hormone replacement. A survey of the first 32 reported patients indicated that genetic males and females were affected with equal frequency, although genetic sex was often inferred from descriptions of gonadal appearance and histology or from buccal smears (Hauffa, et al. 1985). XY genetic male patients are born with female external genitalia due to the absence of testicular testosterone synthesis. Since mitochondria from affected adrenals and gonads fail to convert cholesterol to pregnenolone, the disease was previously thought to be due to a defect in the cholesterol side chain cleavage enzyme, P450scc. However, the involvement of P450scc has been ruled out by molecular genetic analysis of affected individuals (D. Lin et al., *J. Clin. Invest.* **88**, 1955 (1991); Y. Sakai *et al., J. Clin. Endocrinol. Metab.* **79**, 1198 (1994)). We reasoned that the defect could involve the transport of the cholesterol into mitochondria (D. Lin et al., *J. Clin. Invest.* **88**, 1955 (1991); D. Lin et al., *Genomics* **18**, 643 (1993). However, prior to the current elucidation of a molecular defect for lipoid CAH, no specific defect had been found to be associated with this disease.

SUMMARY OF THE INVENTION

Accordingly, it is an object of the present invention to provide a genetic method of diagnosing congenital lipoid congenital hyperplasia in humans.

It is another object of detecting the presence of mutations in a gene responsible for congenital lipoid adrenal hyperplasia in humans for use in genetic counseling.

It is a further object of the invention to provide a method of treating congenital lipoid adrenal hyperplasia in humans by providing a protein that replaces defective proteins in a human with the disease.

It is another object of the invention to provide a method of treating or preventing hypercholesterolemia or other disease caused by aberrant mitochondrial cholesterol metabolism by providing a protein that replaces defective proteins in a human with the disease, provided that the metabolism is governed by enzymes responsive to the protein of the invention.

It is a further object of the invention to provide a promoter for the expression of the protein of the invention, a mutein protein of the invention, or a heterologous gene in a mammalian cell, a transgenic mammal, or expression of a heterologous gene for gene therapy.

These and other objects of the invention as will hereafter become more readily apparent have been accomplished by providing an isolated DNA or RNA molecule, wherein the molecule contains (1) a first sequence consisting of human steroidogenesis acute regulatory protein (hStAR) cDNA, hStAR genomic DNA, hSTAR promoter or a hStAR pseudogene as set forth in Figure 1, Table 6, or Table 7; (2) a second sequence, wherein the second sequence is a subsequence of the first sequence at least 10 nucleotides in length; (3) a third sequence in which at least one nucleotide of the first or second sequence is replaced by a different nucleotide; (4) a fourth sequence in which at least one nucleotide is deleted from or inserted into said first or second sequence; or (5) a fifth sequence complementary to any of the first second, or third sequences; with the provisos that (1) if the molecule is an RNA molecule, U replaces T in the sequence of the molecule, (2) the third sequence is at least 95% identical to the first or second sequence, and (3) the second sequence is not present in mouse StAR cDNA. The invention also provides methods for detecting mutated StAR genes in humans, such mutations having been associated with congenital lipoid adrenal hyperplasia.

BRIEF DESCRIPTION OF THE DRAWINGS

The invention now being generally described, the same will be better understood by reference to the following detailed description of specific embodiments in combination with the figures that form part of this specification, wherein:

Figure 1. The nucleotide and deduced amino acid sequence of the human StAR cDNA (hStAR DNA). The potential sites for protein kinase A and protein kinase C-mediated phosphorylation are noted with single and double underlining, respectively.

Figure 2. Detection of nonsense mutations in patients' StAR cDNAs. **(Top)** RT-PCR products of StAR from normal (NL) human fetal adrenal and testicular RNAs, testicular RNAs of patients 1 and 2, and no RNA control displayed on a 1% ethidium bromide-stained agarose gel. The molecular size markers are *Hind*III-cleaved bacteriophage λ. **(Bottom)** Map of StAR cDNA. RI93→Stop is the substitution of a Stop codon (TGA) for Arg^{193} (CGA) and Q258→Stop is a Stop codon (TAG) for Gln^{258}. The open box represents the coding region of StAR cDNA. The small bars below the map indicate the PCR primers. The sequence of the sense primer S1 was 5'-GCAGCAGCAGCGGCAGCAG-3' (66-84, position in cDNA) and the antisense primer AS1 was 5'-ATGAGCGTGTGTACCAGTGCAG-3' (1016-1037). The PCR program was 94°C. 45 sec; 64°C, 30 sec; 72°C, 60 sec for 30 cycles.

Figure 3. PCR mapping of the StAR gene. **(Top)** Left panel; genomic PCR products amplified with primers S2/AS2 displayed on a 2% ethidium bromide-stained agarose gel. The molecular size markers are *Hae*III-cleaved bacteriophage Φx174. Right panel: genomic PCR products amplified with primers S3/ASI displayed on a 1% agarose gel. The molecular size markers are *Hind*III-cleaved bacteriophage λ. In both gels, genomic DNA was either added as a template in PCR (lane 1) or not added (lane 2). **(Bottom)** Map of the 3' half of the StAR gene. Open boxes represent exons, and numbers labeled at the end of each exon are the corresponding nucleotide position in cDNA sequence (B.J. Clark, J. Wells, S.R. King, D.M. Stocco, *J. Biol. Chem.* **269**, 28314 (1994)). Locations of the various PCR primers and products are shown below the map. The sense primer S2 was 5'GACAAAGTGATGAGTAAAGTG-3' (442-462) and antisense primer AS2 was 5'-TGTGGCCATGCCAGCCAGCA-3' (717-738). The

PCR program using S2/AS2 was 94°C, 45 sec; 58°C, 30 sec; 72°C, 60 sec for 35 cycles. The sense primer S3 was 5'GTGAGCAAAGTCCAGGTGCG-3'. The PCR program using S3/ASI was 94°C, 50 sec; 64°C, 30 sec; 72°C, 90 sec for 35 cycles.

Figure 4. Direct sequencing of PCR products. (A) (Top) Direct PCR sequencing (method of Dynal, Inc., Lake Success, NY) from a normal control, patient 1, and parents of patient 1. Arrows indicate the nucleotide involved in the nonsense mutation: *C* in control, *T* in patient 1, *C* and *T* in both parents. (Bottom) DNA and amino acid sequences. (*B*) Direct PCR sequencing of a normal control, patient 2 and patient 3. Arrows indicate a *C* in the control and a *T* in both patients 2 and 3. In (A), the sense PCR primer (S3) was described in Fig. 2 and the biotinylated antisense primer (AS3) was 5'GGATGCAGTCCACATGCTTGG-3'. The PCR program was 94°C, 45 sec; 64°C, 30 sec; 72°C, 45 sec for 35 cycles. A sense primer, 5'GATACATTCATTACTCAC-3' (613-630) was used for sequencing. In (B), the sense biotinylated primer (S4) was 5'-CCTGGCAGCCTGTTTGTGATAG-3' and the antisense (AS4) primer was 5'-CCTCATGTCATAGCTAATCAGTG-3' (1201-1223). The PCR program was 94°C, 45 sec; 63°C, 30 sec; 72°C, 45 sec for 35 cycles. Antisense primer AS1 was used for direct sequencing.

Figure 5. Mutation causing lipoid CAH. A. DNA sequencing. The normal sequence, shown below, indicates the junction of exons 5 and 6, whereas the patient's sequence shows exon 4 connected to exon 6. B. Intronic mutation. The arrow indicates the T -> A change 11 bp from the junction of intron 4 and exon 5. The sequence is shown below. C. Diagram of the patient's gene and encoded mRNA. The T -> A change in intron 4, indicated as an X, disrupts the splicing of the pre-mRNA, leading to the deletion of exon 5.

Figure 6. Family study. Genomic DNA from a normal control (N1), the patient (Pt), father (Fa) and Mother (Mo) were amplified with primers S2 and AS5,

yielding a 432 bp product. The DNA was electrophoresed uncut (-) or after digestion with *Nco* I (+) compared to a 100 bp size ladder (Markers) and stained with ethidium bromide. A map of the DNA is shown below. Note that the *Nco* I site 74 bp from the 5' end serves as an internal control for complete *Nco* digestion of the DNA.

Figure 7. DNA sequence of the StAR gene promoter. The transcribed sequences from the major transcription start site are indicated in bold letters. Translated sequences are underlined and the amino acids given in single-letter codes. Putative Sp 1 and SF-1 binding sites are indicated. The TATA-like element is boxed. Repetitive sequences are noted with underlines.

Figure 8. PCR diagnosis of lipoid CAH. A: PCR Amplification of genomic DNA from the family of patient 4 using primers HB55 and HB34, followed by *Alu* I digestion. The patient sample yields a 265 bp band, the parents are heterozygous for the 265 and 162 + 103 bp bands and the normal control has only the 162 and 103 bp bands. B: family of patients 5 and 6, amplified with HB55 and HB34, and cut with Tsp45I. The patients have uncut 295 bp fragments, the control has 173 and 122 bp fragments, and the parents are heterozygous. C: genetic variation in alleles carrying R182L is shown in patients 7 and 9. PCR was as in panel B; neither patient's DNA was cleaved by *Tsp*45I; patient 9 also is homozygous for the ΔT593 frame-shift mutation, which destroys an *Ava*II/*Sau*96I site (and also creates *Stu*I site, which yields inconsistent digestions). D: PCR amplification of DNA from patient 10 using primers B2 and AS1, followed by *Fsp*I digestion. The patient's 203 bp fragment is undigested while the control is cleaved to 107 and 96 bp products.

Figure 9. Expression of StAR mRNA in various human tissues. Northern blots containing 2 μg of poly (A)+ RNA isolated from the indicated tissues were purchased from Clontech Laboratories and probed sequentially with StAR and β-actin cDNAs. The autoradiogram in the left hand panel A for StAR was exposed

for 24 h, the right hand panel A autoradiogram for StAR was exposed for 4 h. The blots were both exposed for 2 h for actin (B).

Figure 10. Regulation of StAR mRNA expression in human granulosa cells by cAMP. Primary cultures of human granulosa cells were established in culture for 4 days and then 8-bromo-cAMP (1.5 mM) was added to some dishes (+) for a 24 h period. Results from two separate experiments are presented. Primary cultures of human trophoblast cells were also established in the absence (-) or presence (+) of 1.5 mM 8-bromo-cAMP for 24 h. Total RNA was extracted and subjected to Northern blotting (5 μg RNA/lane) and the blots were probed sequentially with StAR and 28 S rRNA cDNA probes. Autoradiograms were analyzed with an image analysis system (Resource Technology, Nashville, TN) to determine the increase in StAR mRNA in the human granulosa cells relative to 28 S rRNA. The increase was 3-fold in one experiment and 7-fold in the other.

Figure 11. Assignment of the StAR gene to human chromosome 8. Genomic DNA was isolated from a panel of somatic cell hybrids, digested with Hind III and subjected to Southern blotting. The hybrid designation and the human chromosome that predominates, which in some cases is the only human chromosome present in the hybrid, are indicated. A hybridization band corresponding to that found in human genomic DNA was found in a hybrid containing only human chromosome 8. A weaker band was found in hybrid GM 10478, which in addition to chromosome 20 is known to contain a fragment of 8p.

Figure 12.

A. Regional mapping of the StAR gene to 8p by somatic cell hybrid mapping. The chromosome 8 idiogram is modified according to Francke (Francke, U. (1994) Cytogenetics and Cell Genetics 65: 206-219). The right side of the idiogram shows a diagrammatic representation of the portion of human chromosome 8 present in the respective cell lines. The precise localizations of the boundaries of these DNAs on the cytogenetic map of the chromosome are

approximate. The StAR, LPL, SS and CL1 genes were localized by PCR. Presence of a gene is denoted by a '+' and its absence by a '-' symbol. A negative control cell line, CHO-K1, which contains only hamster DNA was also included in these experiments (data not shown). The localization of LPL, SS and CL1 are consistent with previously published data (Wion, K.L., Kirchgessner, T.G., Lusis, A.J., Schotz, M.c., Lawn, R.M. (1987) Science 235: 1638-1641; Fink, T.M., Zimmer, M., Tschopp, J., Etienne, J., Jeene, D.E., Lichter, P. (1993) Genomics 16: 526-528; Schechter, I., Conrad D.G., Hart, I., Berger, R.C., McKenzie, T.L., Bleskan, J., Patterson, D. (1994) Genomics 20: 116-118).

B. YAC FISH localization of the StAR functional gene locus to 8p11.2. YAC DNA was nick translated with biotin dUTP and dCTP and hybridized with metaphase spreads with 1 μg yeast DNA/slide as described in the text. The probe was detected with avidin-FITC (yellow) and chromosomes were counter-stained with propidium iodide (red). The arrow to the left of the idiogram in panel A indicates the FISH location of the A 10 G5 YAC to the 8p11.2 region.

Figure 13. Assignment of StAR pseudogene to human chromosomes 13. PCR analysis of somatic cell hybrid DNA was carried out with primers specific for the StAR pseudogene. The numbers above the lanes in the left hand panel refer to the hybrids analyzed in Figure 7. Hybrid "1" (GM 10880) contains human chromosomes 1 as well as 13 and 14. Hybrid GM 07299A contains human chromosomes X and 1. R370-22A contains human chromosomes 1 and 13. The hybrid designated "13" contains only human chromosome 13. Control designates the cloned pseudogene sequences in pBluescript (Stratagene, La Jolla, CA). The 800 nt StAR pseudogene amplification product is seen only in hybrids containing human chromosome 13.

Figure 14. RNase protection of synthetic RNA assayed by hybridization to radiolabeled probe 1, which consists of the 3' 150 bases of exon 4 (150 bp), exon 5 (185 bp), and 58 bp of vector sequences. The 335 bp protected product

corresponds to exons 4 and 5, the 185 bp product to exon 5,and the group of bands at 150 bp to exon 4.

Figure 15. RNase protection of synthetic RNA assayed by hybridization to radiolabeled probe 2, which consists of the 3' 150 bases of exon 4 (150 bp) the mutated intron 4 (141 bp) exon 5 (185 bp) exon 6 (94 bp), and 68 bp of vector sequences. The probe was hybridized to the following samples before RNase digestion: lane 1 and 11, control single-stranded humal fetal adrenal cDNA; lanes 2 and 3, RNA from COS-1 cells transfected with the mutant (Patient) or Normal StAR minigenes; lane 4, tRNA; lane 7-10 RNA from COS-1 cells transfected with the Patient or Normal StAR minigenes, fractionated into nuclear and cytoplasmic fractions. The sizes and compositions of the principal protected fragments are indicated. Overexposure of the autoradiogram shows that the 638, 570, 471 and 279 base fragments are present in the patient's nuclear RNA (lane 10).

Figure 16. RNase protection of the patient's RNA assayed by hybridization to probe 2 as described in Figure 15. The sizes and compositions of the principal protected fragments are indicated.

DESCRIPTION OF SPECIFIC EMBODIMENTS

The present invention arose in the context of investigations based on the prior indication that placental progesterone synthesis is necessary for the maintenance of pregnancy (J.F. Strauss III et al., in *Endocrinology,* L.J. DeGroot, Ed. (W.B. Saunders, Philadelphia, 1995), vol. 3, pp. 2171-2206). Since pregnancies with a lipoid CAH fetus progress normally to term and the placenta can still produce progesterone (P. Saenger *et al., J: Clin. Endocrinol. Metab.* **80**, 200-205 (1995)), we speculated that the factor sought is required for adrenal and gonadal, but not placental, steroidogenesis. A recently described 30-kDa phosphorylated protein is believed to mediate the rapid and cycloheximide-sensitive response of steroidogenesis to tropic stimulation (D.M. Stocco and T.C. Sodeman, *J. Biol. Chem. 266,* 19739 (1991); L.F. Epstein and N.R. Orme-Johnson, *J. Biol.*

Chem. **266**, 19739 (1991); D.M. Stocco and M. Ascoli, *Endocrinology* **132**, 959 (1993)). This protein, termed steroidogenesis acute regulatory protein (StAR), was purified from MA-10 murine Leydig tumor cells. The cloning of StAR cDNA from mouse was previously described in the scientific literature (B.J. Clark, J. Wells, S.R. King, D.M. Stocco, *J. Biol. Chem.* **269**, 28314 (1994)). In order to determine whether our hypothesis was correct, i.e., that a genetic defect in this protein could be responsible for lipoid CAH, we cloned human StAR cDNA. The nucleotide and deduced amino acid sequence of the human StAR cDNA is provided in Figure 4. Mutations are described using the nucleotide numbering system in Figure 4 and the affected amino acid is described by a numbering system wherein the methionine encoded by nucleotides 127-129 is the first amino acid of the StAR protein.

Transient expression of mouse StAR cDNA in MA-10 cells and COS-1 cells results in enhanced steroidogenesis (B.J. Clark, J. Wells, S.R. King, D.M. Stocco, *J. Biol. Chem.* **269**, 28314 (1994)). We discovered similar properties for human StAR cDNA and further found that StAR mRNA is abundant in adrenal and gonad tissue, but not in placenta. Thus StAR appeared to be a good candidate for the factor involved in lipoid CAH. This prompted us to examine the StAR gene in nineteen unrelated patients. Patient 1, of Caucasian ancestry, has not been reported previously by others skilled in the art, Patient 2, an ethnic Korean, and Patient 3, an ethnic Japanese, were previously described, but not with regard to the relevance of the StAR gene (Patient 3: B.P. Hauffa et al., *Clin. Endocrinol.* **23**, 481 (1985).; Patients 2 and 3: D. Lin et al., *J. Clin. Invest.* **88**, 1955 (1991); Patient 2: D. Lin et al., *Genomics* **18**, 643 (1993).

We generated StAR cDNA from Patients 1 and 2 by reverse transcription-polymerase chain reaction (RT-PCR) using testicular mRNA as template. When PCR primers from the 5' and 3' untranslated regions were used, the principal product was StAR cDNA, but there were related species that contained a large number of sequence differences. This lead to the discovery of a StAR pseudogene reported in the examples below. Using a sequence termed S1 in the 5' untranslated region that distinguishes authentic StAR from its pseudogene, we amplified the

974-bp StAR cDNA in normal controls and in two patients (Figure 2). These RT-PCR products were subcloned into pCRII vectors and sequenced. All patient clones from independent RT-PCR reactions were identical to the wild type sequence except for a *C* to *T* transition in codon 193 (Arg) in Patient 1 and a *C* to *T* transition in codon 258 (Gln) in Patient 2. These generated premature stop codons, leading to mutant proteins lacking 93 or 28 amino acid residues, respectively, from the C-terminus.

To confirm the identity of these mutations, we analyzed StAR genes from genomic DNA of our patients. Since the structure of the StAR gene was unknown, we first used PCR to obtain a genomic clone containing the exons harboring the mutations. This was done by using various combinations of sense and antisense primers derived from the cDNA sequences to amplify normal genomic DNA. As shown in Figure 3, the primer pair S2/AS2 yielded two specific products of 437 bp and 290 bp. The sequence of the 437-bp fragment matches the cDNA sequence at both ends perfectly and contains a 141-bp intron in the middle, thus deriving from the StAR gene. The 290-bp fragment was from the StAR pseudogene, lacking the intron. Subsequently, an intronic primer termed S3 was used with primer AS1 for PCR, which yielded a 2.1-kb product (Figure 3). Mapping and DNA sequencing of this fragment revealed that the sequences of the exons match perfectly with the cDNA and all intron/exon boundaries strictly follow the GT/AG rule. Thus the 2.1 kb fragment represents the 3' half of the StAR gene. The sequence information obtained from the 2.1-kb clone enabled us to make intronic primers to PCR-amplify the exons (Figure 3).

The presence of the nonsense mutations in codons 193 and 258 was confirmed by directly sequencing PCR products of the genomic DNA. As shown in Figure 4A, Patient 1 has a *C* to *T* transition at codon 193, whereas her father or mother have *C* and *T* at codon 193 (one on each of their two chromosomes). Therefore, we conclude that Patient 1 is homozygous for the $Arg^{193}\rightarrow Stop$ mutation, and both of her parents are carriers for this mutation. Similarly, Patient 2 is homozygous for the $Gln^{258}\rightarrow Stop$ mutation (Figure 4B). As expected, the mother of Patient 2 was heterozygous for this mutation, while a normal sibling had no

mutation. Patient 4 was a sibling also afflicted with lipoid CAH and thus is homozygous for the same allele as Patient 2. In addition, Patient 3 was homozygous for the same mutation as Patient 2 (Figure 4B); her mother is also a carrier. Since patient 2 is an ethnic Korean and patient 3 is an ethnic Japanese, this finding suggests a common origin for this mutation in these two ethnic groups.

To prove that these premature stop codons in StAR cause functional alterations, we analyzed the expressed wild-type and mutant proteins for their ability to enhance steroidogenesis. Using lipofectamine, nonsteroidogenic COS-I monkey kidney cells were transfected with pSPORT (Vector) or with pSPORT expressing normal human StAR or the mutant StAR from Patients 1 and 2 (or 3). The cells were co-transfected with either vectors expressing bovine P450scc and bovine adrenodoxin (both provided by Dr. Michael Waterman, Vanderbilt University), or a pECE vector expressing a fusion protein termed F2, consisting of the human cholesterol side-chain cleavage system: H_2N-P450scc-Adrenodoxin Reductase-Adrenodoxin-COOH (J.A. Harikrishna et al., *DNA Cell Biol.* **12**, 371 (1993)). The substrate was either the cellular and serum cholesterol (chol) or added 5 μg/ml 20α-hydroxycholesterol (20α). After 48 h of incubation, the medium was collected and assayed for pregnenolone by immunoassays. The results are shown in Table 1. Co-expression of StAR with the cholesterol side chain cleavage system resulted in an approximately eight-fold increase in pregnenolone production when cholesterol was used as a substrate. Both mutant StAR proteins are inactive, indicating that each of the two nonsense mutations causes lipoid CAH. Unlike cholesterol, 20α-hydroxycholesterol can readily diffuse into the mitochondria and thereby bypasses the mitochondrial cholesterol transport system (M.E. Toaff et al., *Endocrinology, III* 1785 (1982)). With 20α-hydroxycholesterol as a substrate, there are no significant differences in pregnenolone production between normal StAR and mutant StARs. The differential effects of StAR on utilization of cholesterol and 20α-hydroxycholesterol strongly suggest that StAR mediates the transport of cholesterol into mitochondria.

Table 1

Loss of StAR activity due to nonsense mutations

Pregnenolone Production (ng/dish)

Co-transfection	scc/Adx		F2	
	chol	20α	chol	20α
Vector	20 ± 1	158 ± 21	17 ± 3	60 ± 7
StAR	175 ± 19	138 ± 15	131 ± 23	60 ± 11
Patient 1	19 ± 2	99 ± 23	18 ± 5	56 ± 7
Patient 2 or 3	25 ± 4	168 ± 35	22 ± 4	75 ± 7

Values are the means ± standard deviations from four separate transfections.

StAR is synthesized as a 285 amino acid protein with a mitochondrial targeting sequence of 25 residues, which is cleaved from the N-terminus following transport into mitochondria. The precursor and mature StARs have half-lives in the range of minutes and hours, respectively. Digital videoscanning of immunoblots (Clark et al. 1994) revealed that about 70% of StAR in COS-1 cells transfected with wildtype plasmid was in the mature form. However, no mature form was seen for mutant protein from Patient 1 and about 10% was processed for Patient 2 (not shown), suggesting a possible mechanism for the loss of activity.

We have also identified an unusual intronic mutation in the StAR gene that results in an mRNA splicing error, thus causing lipoid CAH (Tee, M.K. et al. *Hum. Mol. Genet.* 4, 2299-2305 (1995c)). Initial analysis of the StAR cDNA prepared from the testicular tissue of Patient 5 showed that it was smaller than the 974 bp band obtained from normal tissue, indicating a major structural alteration in the patient's StAR mRNA. Sequencing of the cloned cDNA indicated that all of exon 5 was absent from the patient's cDNA (Figure 5). This resulted in a shorter cDNA of only 789 bp that lacked the 185 bp corresponding to exon 5.

Deletion of exon 5 suggested that there might be an mRNA splicing error, possibly in intron 4 just upstream from exon 5. Therefore, we used the primers S3, which lies in intron 4, and AS2, which lies in exon 5, for PCR amplification of genomic DNA from the patient. This DNA was subcloned and sequenced using the primer S3, which lies in intron 4. Only a single nucleotide change was found in intron 4, a T→A transversion 11 bp from the junction of intron 4 and exon 5 (Figure 5).

The T→A transversion disrupts an *Nco* I site (CCATGG→CCAAGG), permitting confirmation of the Mendelian segregation of this base change. Analysis of an *Nco* I restriction digest of amplified genomic DNA established that Patient 5 is homozygous for the T→A transversion (Figure 6). RNase protection assays performed on RNA obtained from cells transfected with T→A transversion and the testicular RNA from Patient 5 established that the T→A transversion causes several forms of disordered mRNA splicing (Example 9).

To determine if the encoded truncated protein was active, we cloned truncated StAR cDNA into the expression vector pSV-SPORT-1 and used this to transfect non-steroidogenic COS-1 cells that were co-transfected with a vector expressing a fusion protein of the human P450scc system, termed F2. The encoded F2 protein is H_2N-P450scc-Adrenodoxin Reductase-Adrenodoxin-COOH, which is a fully active enzyme that eliminates variations in P450scc activity attributable to variations in the molar ratios of the three components (Harikrishna et al.). The cells transfected with F2 and normal StAR supported efficient conversion of cholesterol to pregnenolone, but co-transfection of F2 with the vector expressing the truncated StAR resulted in no more conversion of cholesterol to pregnenolone than was seen in control cells transfected with the pSV-SPORT-1 vector alone (Table 2). Northern blotting of the RNA from the transfected cells sowed that the vectors expressing the normal and mutant StARs expressed equivalent amounts of StAR mRNAs of the predicted sizes, indicating the mRNA was stable. However, Western blotting with antiserum to mouse StAR (generously provided by Dr. D. Stocco, Texas Tech U.), which detected normal StAR protein expressed by pSV-SPORT-1, did not detect any truncated StAR encoded by the

mutant. Thus the truncated StAR mRNA did not encode a stable, functional protein.

Table 2

Activity of normal and mutant StAR expression vectors.

Pregnenolone Production (ng/dish)

	Experiment 1		Experiment 2	
	Cholesterol	20α	Cholesterol	20α
Vector	21 ± 4	57 ± 7	17 ± 3	60 ± 7
pStAR	117 ± 13	48 ± 6	131 ± 23	60 ± 11
Patient 5	29 ± 1	62 ± 5	20 ± 7	70 ± 8

Values are the means ± SD from three dishes in experiment #1 and 4 dishes in experiment #2.

Having found two nonsense mutations in four affected patients (Patients 1, 2, 3, and 4) and a splicing error in a fifth patient (Tee, et al.) and having demonstrated that StAR mutations could cause lipoid CAH, we examined the StAR gene in fourteen additional patients from various ethnic groups to determine whether all patients with the lipoid CAH phenotype have StAR mutations.

We obtained genomic DNA from fourteen patients with lipoid CAH (Table 3). Exons 1-4 were PCR-amplified individually and subjected to automated sequencing, revealing only one mutation, a nucleotide insertion and frame-shift on one allele in Patient 7. Exons 5-7 were amplified as a single 2.1 kb fragment, which was cloned and manually sequenced in its entirety, including introns, for all patients. Identified mutations were confirmed by direct sequencing of the individual PCR-amplified exon of the affected individual and the parents and siblings whenever possible. Many of the mutations could also be confirmed by RFLP analysis of PCR-amplified DNA (Table 4). Table 3 summarizes the findings of these 14 patients and our 5 previously described patients. Patients 11 and 12

and Patients 2 and 4 are sib pairs and Patient 9 is from a known consanguineous marriage, thus these 19 patients represent 33 distinct alleles.

Table 3

Lipoid CAH Patients

Patient No.	Ethnicity or Nationality	Nucleotide Mutation[a]	Protein Mutation[a]	Exon	Affected Alleles[b]	Karyotype
1	Caucasian	C703T	R193X	5	2	XY
2	Korean	C898T	Q258X	7	2[c]	XY
3	Japanese	C898T	Q258X	7	2	XY
4	Korean	C898T	Q258X	7	2	XX
5	Vietnamese	T→A @-11[14]/[ES]	Frame	5	2	XY
6	Japanese	C898T	Q258X	7	[c]	XY
7	Japanese	247/InsG/248 C898T	Frame Q258X	2 7	1 1	XY
8	Japanese	G631A C898T	E169K Q258X	5 7	1 1	XY
9	Palestinian	A632G	E169G	5	2	XY
10	Palestinian	G671T	R182L	5	2	XX
11	Palestinian	G671T	R182L	5	2[c]	XY
12	Palestinian	G671T	R182L	5	[d]	XY
13	Palestinian	ΔC650 G671T	Frame R182L	5 5	2 1	XY
14	Palestinian	ΔT593 G671T	Frame R182L	5 5	2 2	XY
15	Mexican	Δ940-942	ΔR272	7	2	XY
16	Greek	947/InsA/948	Frame	7	2	XX
17	British Caucasian	Del/Ins	Truncate	5		XY
18	Canadian Caucasian		L275P A218V	7 6	1 1	XY

Footnotes to Table 3

[a] Nucleotide and amino acid numbers are given according to the cDNA sequence (Sugawara, et al. *Proc. Natl. Acad. Sci. USA* **92**, 4778-4782 (1995a)).

[b] Compound heterozygotes have two entries per patient (one allele), homozygotes have a single entry (two alleles). Only Patient 10 is from a known consanguineous union.

[c] Patients 2 and 4 are siblings.

[d] Patients 11 and 12 are siblings.

Table 4

Molecular Diagnosis of Lipoid CAH

				Oligonucleotides				
	Nucleotide Mutation	Protein Mutation	Exon	Sense	Antisense	Restriction Enzyme	Normal Fragments	Mutant Fragments
1	C898T	Q258X	7	S4	AS4	*Eco*RII	115, 324	439
2	A632G	E169G	5	Ex5S	Ex5AS	*Alu*I	162, 103, 30	265, 30
3	G671T	R182L	5	Ex5S	Ex5AS	*Tsp*45I	173, 122	295
4	ΔT593	Frame	5	Ex5S	Ex5AS	*Sau*96I or *Ava*II	204, 91	295
5	Δ940-942	ΔR272	7	B2	AS1	*Fsp*I	107, 96	203
6	947/InsA/948	Frame	7	B2	AS1	*Hha*I	108, 91, 6	108, 98
7	C703T	R193X	5	Ex5S	Ex5AS	*Hae*II	295	197, 98
8	T→A @-11[M/Es]	Frame	5	S2 Ex5S	AS5 Ex5AS	*Nco*I	207, 151, 74	358, 74
9	ΔC650	Frame	5	S3	Ex5AS	*Hae*III	99, 75, 68, 62	99, 89, 75, 68, 62

Q258X accounts for most lipoid CAH in Japan. The Q258X mutation was homozygous in Patients 2, 3, and 4 (i.e. 4 of 4 unique alleles) and was present in 4 of 6 Japanese alleles (Patients 6-8), for a total of 8 of 10 affected Japanese/Korean alleles. The other mutations found in Japanese patients have not been found in other individuals and may represent new mutations, whereas Q258X

mutation appears to represent a founder effect. Thus, this mutation accounts for a preponderance of the affected alleles in Japan, where lipoid CAH is common (Hauffa, et al.; Fukami, et al. *Clin. Pediatr. Endocrinol.* **4**, 39-46 (1995); Matsuo, et al. *Horm. Res.* **41 (Suppl)**, 106 (1994)). The Q258X mutation is easily identified by amplifying genomic DNA with primers S4 and AS4 (Lin, et al. *Science* **267**, 1828-1831 (1995)) followed by digestion with *Eco* RII, as the responsible C→T mutation (underlined) destroys *Eco* RII site CCAGG (Table 4).

R182L is a common mutation in Palestinian Arabs. Lipoid CAH has previously been described as occurring disproportionately more commonly in two populations; the Japanese and the German Swiss (Hauffa, et al.), but this disorder has not been described among Arabs. However, six of our patients were of Palestinian Arab ancestry. Patients 10 and 11 were siblings, and Patient 9 was the result of a known consanguineous marriage, so that these 6 patients represented 9 unique alleles. Seven of these nine alleles (78%) bore the mutation R182L. These patients came from Jordan, Israel, Kuwait, and Denmark and hence appeared to be unrelated. Identification of intronic polymorphisms and other mutations within the StAR gene established that the 7 affected alleles were not wholly identical. No other mutations were found in sibling Patients 10 and 11, or in Patient 12. Patient 13 who was heterozygous for R182L, was also homozygous for the frame-shift mutation of ΔC650. Patient 14 was doubly homozygous for the frame-shift mutation ΔT593 and for R182L. Thus the R182L mutation was found in various sequence contexts, and was strongly associated with the Arab Palestinian population. The R182L mutation is easily identified by amplifying genomic DNA with primers Ex5S and Ex5AS followed by digestion with *Tsp*45I (Table 3).

StAR mutations cluster in exons 5-7. We examined 5 additional patients who were neither Japanese or Arab. Patient 15, a Mexican of Native American ancestry was homozygous for 3 bp deletion that deleted R272 but otherwise left the StAR protein intact. Patient 16, from Greece, was homozygous for a frame-shift mutation. Patient 17, a Caucasian from Britain was homozygous for insertion of a foreign DNA segment beginning in exon 5. Although the StAR gene in these three cases was homozygous for each mutation, no history of consanguinity could

be elicited from these families. Patient 18, a Caucasian from Canada, was a compound heterozygote for L275P and A218V. Among the 33 alleles investigated in Table 2, we have found a total of 14 mutations, all but one of which affected exons 5, 6, or 7.

To prove that the identified mutations caused the patients' lipoid CAH, we assayed the conversion of cholesterol to pregnenolone by conditioned media of non-steroidogenic COS-1 cells transfected with the various StAR mutations as described in Example 3.

StAR function is not associated with StAR processing. It has been suggested that the active form of the StAR protein is the precursor molecule, rather than the cleaved intramitochondrial form, and that the StAR precursor acts to stimulate steroidogenesis by forming contact sites between the outer and inner membranes as it enters the mitochondria (Clark, et al. *J. Biol. Chem.* **269**, 28314-28322 (1994); Stocco, et al. *Cellular and Molecular Regulation of Testicular Cells* (eds Desjardins, C.) (Springer-Verlag, New York (1996)). However deletion of as few as 28 carboxy-terminal amino acids of StAR as found in the common Japanese mutation Q258X, deletes all activity (Lin, et al. 1995), and the data in Table 4 show that deletion of R272 or replacement of amino acids at positions 169, 182, 218, and 275 ablate all activity. With the exception of an exon 2 frame-shift in Patient 7, all of the mutations we have found lie in exons 5-7. Thus either exons 1-4 are less prone to spontaneous mutations, or mis-sense mutations in this region are phenotypically silent; we favor the latter explanation.

Mutations in the transcribed region of the human StAR gene cause lipoid CAH if they result in alterations of the StAR protein or alterations in control of its production. Until the molecular basis of lipoid CAH was determined, the disorder had to be considered a syndrome, potentially a group of different genetic diseases with a single common phenotype. The identification of mutations in the gene for the StAR protein in lipoid CAH patients from a wide variety of ethnic and genetic backgrounds now clearly establishes that mutations in this gene is responsible for the overwhelming majority, if not all patients with lipoid CAH. The failure to detect StAR mutations in one CAH patient (Patient 19) may be due to a technical

problem in detection of mutations, a promoter mutation, an uninvestigated upstream splicing mutation, a recurring sequencing error, or Patient 19 may have a mutation in another gene giving rise to an indistinguishable phenotype. The demonstration that the placenta of the affected fetus continues to produce steroids normally in CAH (Saenger, et al. *J. Clin. Endocrinol. Metab.* **80**, 200-205 (1995)), and the fetal requirement for progesterone throughout gestation, make it most unlikely that any of the factors previously considered in lipoid CAH (P450scc, Adx, AdRed, SCP2, SAP endozepine, PBR) could account for the disease in Patient 19.

Genotypic sex in lipoid CAH is most commonly male, a finding inconsistent with early studies that showed that genetic males and females were affected with equal frequency (at that time, however, genetic sex was often inferred from descriptions of gonadal appearance and histology or from buccal smears (Hauffa, et al.)). Table 2 shows that only 3 of 19 patients (15.8%) in our series are 46,XX, and a preliminary report from Japan found only sixteen 46,XX patients among 63 (25.4%) lipoid CAH patients with established karyotypes (Matsuo, et al. 1994). This suggests either that a large portion of affected 46,XX fetuses are lost in early pregnancy, or that affected 23X sperm are less likely to fertilize an egg, or are produced with lower frequency than are affected 23Y sperm. Some preliminary data suggest that steroidogenesis may occur in spermatogonia, possible supporting the hypothesis that the sex bias occurs before fertilization. However, an ascertainment bias of unknown nature cannot be ruled out at the present time.

Lipoid CAH is the only known inborn disorder of steroid hormone synthesis not caused by a defective steroidogenic enzyme. The identification of mutant StARs in lipoid CAH now permits prenatal molecular diagnosis for this devastating disease. Lipoid CAH due to nonfunctional StARs is comparable to the effect of a StAR gene knockout, demonstrating that StAR is indispensable for adrenal and gonadal steroidogenesis. Thus, StAR is the first protein identified that plays an essential role for cholesterol access to P450scc. The sparing of fetuses with lipoid CAH as a result of the presence of normal placental steroidogenesis and the absence of StAR expression in placenta (as we discovered) and in other

steroidogenic tissues, such as brain (P. Robel and E.E. Baulieu, *Trends Endocrinol. Metab.* **5**, 1 (1994); S.H. Mellon, *J. Clin. Endocrinol. Metab.* 78, 1003 (1994)), suggests that different mechanisms may exist to facilitate cholesterol transport into mitochondria in these tissues. This demonstration of the critical role of StAR in lipoid CAH provides the first genetic evidence for the hypothesis that StAR is the long-sought molecule that mediates the acute tropic regulation of steroid hormone synthesis (D.M. Stocco and T.C. Sodeman, *J. Biol. Chem. 266,* 19739 (1991); L.F. Epstein and N.R. Orme-Johnson, *J. Biol. Chem.* **266**, 19739 (1991); D.M. Stocco and M. Ascoli, *Endocrinology* **132**, 959 (1993)).

However, StAR's actions are not specific to steroidogenesis and StAR can also stimulate mitochondrial 27-hydroxylase activity. Mitochondrial cholesterol 27-hydroxylase (P450c27) catalyzes the formation of 27-hydroxycholesterol and the C_{27} acid, 3β-hydroxy-5-cholestenoic acid (Andersson et al., *J. Biol. Chem.* **264**, 8222-8229 (1989); Su et al., *DNA Cell. Bio.* **9**, 657-665 (1990)). Expression of StAR in COS-1 cells cotransfected with P450c27 and adrenodoxin resulted in a more than 6-fold increase ($p < 0.005$ for comparison of group 3 vs group 4) in the production of 3β-hydroxy-5-cholestenoic acid as shown in Example 10. This StAR-mediated increase in cholesterol metabolism by P450c27 is on the same order of magnitude as we observed for StAR stimulation of cholesterol side-chain cleavage (Sugawara et al., *Proc. Natl. Acad. Sci. USA* **92**, 4778-4782 (1995)). These observations demonstrate that StAR is capable of enhancing mitochondrial cholesterol metabolism by enzymes other than the steroidogenic cytochrome P450scc.

P450c27 is found in a number of tissues, including the ovary which expresses StAR (Andersson et al.; Su et al.). Thus, StAR could contribute to the metabolism of cholesterol to 27-hydroxycholesterol and the C_{27} acid in steroidogenic cells. The hydroxysterols so-formed could have roles in governing cellular cholesterol homeostasis (Rennert et al., *Endocrinology* **127**, 738-746 (1990)).

P450c27 is highly expressed in the liver where it play a role in bile acid synthesis (Andersson et al.). Because the StAR gene is not expressed in the liver

(Sugawara et al., PNAS (1995)), other factors or processes must govern cholesterol access to the hepatic 27-hydroxylase. The StAR gene can be introduced into the liver to enhance bile acid formation through 27-hydroxylase pathway, and hence promote cholesterol disposal.

Because mutations arising outside of the coding region can also cause lipoid CAH as was observed, for example, in Patient 5, we further characterized the genomic DNA and promoter region of the StAR gene. It is possible that in Patient 19 lipoid CAH is caused by a mutation in the promoter that interferes with transcription. Mutations arising in the 5' and 3' untranslated regions might likewise interfere with transcription or translation by, for example, *e.g.* altering mRNA stability, or affecting translation initiation.

The StAR gene sequence comprises seven exons and the exon-intron junctions (Table 5) all obey the GT/AG rule (Mount, 1982).

Table 5

Exon-Intron Junctions of the Human StAR Gene

			sequence at the exon-intron junctions				
			5'-splice donor			3'-splice acceptor	
exon no.	nucleotide position	length (bp)			intron size		
I	1-203	203	CATGAAGG	gtgagcgctgcgggaaggaggcga	2 kb	aacaagggttanccctctgcag	GGCTGAGG
II	204-318	115	TCTACTCG	gtaagtgctgaggcttctgggctc	0.3 kb	gtctctcctcggctgtgtatccag	GTTCTCGG
III	319-445	126	GTCAGCAG	gtaagtgtcggggagaagccctgtg	1.8 kb	tctgggggctccmctctgacag	GACAATGG
IV	446-605	159	AGATCAAG	gtgagcaaagtccaggtgcgggtg	0.14 kb	nctggttccccatggcctggtag	CTCCTGCA
V	606-791	185	GTCATCAG	gtaatacgggcagcaggctccaaac	0.65 kb	gacntgacntgctccamgccag	GGCGGAGC
VI	792-886	94	ACCTCAAG	gtgaagggcatgggaggggggacct	0.85 kb	aaantctcctacctcctactgcag	GGGTGGCT
VII	897-1634	747					

We performed extensive sequence analysis on the original genomic clone disclosed in US Patent Application Number 08/410,540 from which this application is a continuation-in-part. The sequence for the coding regions and intron 4 as disclosed in the parent application was accurate, but upon further characterization of the DNA molecule of the invention, sequencing errors were identified and

corrected in the sequence 5' of the coding region. The corrected sequence is provided in Table 6.

Table 6

Genomic DNA Sequence of Steroidogenesis Acute Regulatory Protein Gene

AGCTTTCTGCACATACCAAGACCCCCAGCCCAGCTCACTCAGACAAAG
CTACTGGCGGGAAAGTGTGAGGAAGGGTGTGGGCGTGGCCCAGGCCCT
CCTCTTCTCTCTGCCGTATACTGATAGGGCTGCCCCGCACCCCCCCCCG
CCCCCCCGCGACTCAGCCACGAGAGGTCATCCTTGCTCCAGCACAAGA
CCCCTAAGAACCCTCACTCGGAACAGGACTTTGGGAAAGGTGGTTTTC
TATAAATAGATGAGTAAATAAATTGACAGTTGATATACCAAGCGTCCT
GGGGCCGCAGGAGGAACTGTGTACAGATGGCTTGAAGGCCAGAGGCTT
GGGTTGTTGTACTGGCCCTTCCACTGGCCAGCTGTTTGACCTTGAACAA
TCAAGTTCCACTCTGTGGACTTCAGGGTCCTCACCCAGAAGAAGAGCA
GCCATATGGTCTCTACTGCCTGGTAAACACCCTGGCTCACTCTCGCGA
GATGGTGGTTCTCCAAGTGTAGTGTGTAGTCCACACAACACCTGCATT
GCAACCACTGGGTATTTATTTATTTATTTAATTTATTTATTTAATTTATG
ATGGAGTCTCACTCTGTCGCCCAGGATGGAGTGCAGTGGCACGATCTT
GGCTTACTGCAACCTCTGCCTCCTGGGTTCAAGTGATTCTCATGCCTCA
GCCTCCCGAGTAGCTGGGACTACAGGTGCCTGCCACATCACCCGGCTA
ATTTTTTGTATTTTTAGTAGAGATGAGGTTTCACCATGTTGGCCATTCT
GGTCTCGGACGCCTGACCTCAAGTGATCTTCCCACCTCGGCCTCCCCA
AGTTCTGGGGTTACAGGCGTAAACCACCGCCCCTGGCCAAGGGGAGGT
TTTTTCCTTTCCTTTTTTCCTTTTTCCTTTCCTTTCCCTCCTTTTTTTTTT
TTTTTTTTTTTTTTTTTTTAACACAGGTTTCTGAGCCTCAATTCCAGATC
AGCTGAGCCTGGAGTTTCTGAAGACAAGGGCTAGAAATCTGCACTTTA
AAGTCTTGAAAACCACTGTGTGCCTTCATCTAAGCTGCCCCTGCTTCTC
TCCCCTCCATCCCTCGCCTGGCCCTGTCCTCCCTACTCTCCCCTGCACC
CTCCCCCGCCCCAAGCTCCCCACAAACGGCCAAAGCAGCAGTGTGAGG
CAATCGCTCTATCCTTGACCCCTTCCTTTGCACAGTGAGTGATGGCGTT
TTTATCTCCTGATGATGATGCACAGCCTTCAGCGGGGGACATTTAAGA

CGCAGAACACCAGGTCCAGGCTGCAGCTGCGGGACTCAGAGGCGACT
CAGAGGCGAAGCTTGAGGGGCTCAGAAGGACGAAGAACCACCCTTGA
GAGAAGAGGCAGCAGCAGCGGCGGCAGCAGCAGCGGCAGCGACCCAC
CACTGCCACATTTGCCAGGAAACAATGCTGCTAGCGACATTCAAGCTG
TGCGCTGGGAGCTCCTACAGACACATGCGCAACATGAAGGGTGAGCGC
TGCGGGAAGGAGGCGATGAGGGGTTGGCCAGCTCTCAGCGGATGAGG
CTCAGGCCACCCAATTCTGATCCTAGTTGTGCCTCTTACTGGGTGAACC
TGGGCAAGTTTCTTCCCTTCTTGAATCTCAGTTTTCCCCTCGGAAGGGA
GCACTACCATGGGAGNTGAGGTNCTGGCTCTAGTTCAGGTCCCTGCTA
GAATACTGTGTTNTNNTGAGCAAGNCACATCCCTCTCCACNCCCACTT
ACTCATTTGAGANTANATGANGGGGTGGNGTGGGCCATCTCTAAGGGG
CTTNGCCAGCTCCTAGACAANGGNTATTCCCTTCTCCAGGGCTGAGGC
AACAGGCTGTGATGGCCATCAGCCAGGAGCTGAACCGGAGGGCCCTG
GGGGGCCCCACCCCTAGCACGTGGATTAACCAGGTTCGGCGGCGGAGC
TCTCTACTCGGTAAGTGCTGAGGCTTCTGGGCTCCTGGTGCTGCTGGCA
GGAGGTTCCCTGGAGGGTGATGTGGTGCATGTGGCTTTGGCTCCCCTC
CTGCCATTCCTTCATTTTGAGAGGACGTCCCCAGCCTAGAGTTCCTCAA
GGCCAGATCCCTCTCTGGTCACCTGGGGCGGCTGTGATTAACTCGACC
AGCAGGCTGGCCCCTATGGCTTTAGTCCGGGCTCTTCAGAGCAATGAG
CAGACCCAGAGCTCCAGGGATGAGAGCTGGTGGAGGCTGGGAGAAGA
AGGAAGCTCTGTCTCTCCTCGGATGTGTATCCAGGTTCTCGGCTGGAA
GAGACTCTCTACAGTGACCAGGAGCTGGCCTATCTCCAGCAGGGGGAG
GAGGCCATGCAGAAGGCCTTGGGCATCCTTAGCAACCAAGAGGGCTGG
AAGAAGGAGAGTCAGCAGGTAAGTGTCGGGGAGAAGCCTGTGGTTCCT
CCATATGCCCGGCCAAGAATATTTTTGTCTAACCACCTTCTGGGGGCTC
CTTTCTCTGACAGGACAATGGGGACAAAGTGATGAGTAAAGTGGTCCC
AGATGTGGGCAAGGTGTTCCGGCTGGAGGTCGTGGTGGACCAGCCCAT
GGAGAGGCTCTATGAAGAGCTCGTGGAGCGCATGGAAGCAATGGGGG
AGTGGAACCCCAATGTCAAGGAGATCAAGGTGAGCAAAGTCCAGGTG
CGGGTGGCAGGGGCCCAGGAGAGCCCAGTGTGAATGCTGTATCAAAG
AGAGGACCCCTAGCTGTGGGGGGTGCTTAGCCCAACACAGGCTGAGTC

GTGATTCTGGTTCCCCATGGCCTGGTAGGTCCTGCAGAAGATCGGAAA
AGATACATTCATTACTCACGAGCTGGCTGCCGAGGCAGCAGGAAACCT
GGTGGGGCCCCGTGACTTTGTGAGCGTGCGCTGTGCCAAGCGCCGAGG
CTCCACCTGTGTGCTGGCTGGCATGGACACAGACTTCGGGAACATGCC
TGAGCAGAAGGGTGTCATCAGGTAATACGGGCAGCAGGCTCCAAACCC
CCCNAGGANTCCCCACTTTCCNCCTNACCTNACNTTCCCCAATTTCCA
GGGCGGAGCACGGTCCCACTTGCATGGTGCTTCACCCGTTGGCTGGAA
GTCCCTCTAAGACCAAACTTACGTGGCTACTCAGCATCGACCTCAAGG
TGAAGGGCATGGGAGGGGGACCTGGAAGGCAGGTTATGNGANAGGGT
GCAGANTCAANCNTGGTGCATAGNCCACAAGATGAGCACATTCTCCTA
CCACCTACTGAAGGGGTGGCTGCCCAAGAGCATCATCAACCAGGTCCT
GTCCCAGACCCAGGTGGATTTTGCCAACCACCTGCGCAAGCGCCTGGA
GTCCCACCCTGCCTCTGAAGCCAGGTGTTGAAGACCAGCCTGCTGTTC
CCAACTGTGCCCAGCTGCACTGGTACACACGCTCATCAGGAGAATCCC
TACTGGAAGCCTGCAAGTCTAAGATCTCCATCTGGTGACAGTGGGATG
GGTGGGGTTCGTGTTTAGAGTATGACACTAGGATTCAGATTGGTGAAG
TTTTTAGTACCAAGAAAACAGGGATGAGGCTCTTGGATTAAAAGGTAA
CTTCATTCACTGATTAGCTATGACATGAGGGTTCAGGCCCCTAAAATA
ATTGTAAAACTTTTTTTCTGGGCCCTTATGTACCCACCTAAAACCATCT
TTAAAATGCTAGTGGCTGATATGGGTGTGGGGGATGCTAACCACAGGG
CCTGAGAAGTCTTGCTTTATGGGCTCAAGAATGCCATGCGCTGGCAGT
ACATGTGCACAAAGCAGAATCTCAGAGGGTCTCCTGCAGCCCTCTGCT
CCTCCCGGCCGCTGCACAGCAACACCACAGAACAAGCAGCACCCCAC
AGTGGGTGCCTTCCAGAAATATAGTCCAAGCTTTCTCTGTGGAAAAAG
ACAAAACTCATTAGTAGACATGTTTCCCTATTGCTTTCATAGGCACCAG
TCAGAATAAAGAATCATAATTCACACCAAACATCAGTCTTTGTTTTAAT
ATTGTACTTGTTAAAAAAATCTATGCAGCTGGGTGCAGTGGCTCACGC
CTGTAATCCCAGCATTTTGGGAGGCTGAGGTAGGCGGATCGAGTCGAC
TCCCTTTAGTGAGGGTTAATTGAGCTCCACCGCGGTGGCGGCCGCTCT
AGAACTAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATC
GATACCGTCGACCTCGAGGGGGGGCCCGGTACCCGGA

The major transcription start site of the StAR gene, determined by RNase protection analysis and confirmed by primer extension analysis, lies 154 bp in front of the translation start site (Figure 7). A minor site was identified about 35 bp further upstream of the major site. The DNA sequence upstream from the major transcription start site contains a TATA-like element (TTTAA) at —24 to —20 bp. Several putative Sp1 binding sites are also present in the promoter region as well as repeats of the sequence ATTT and $(T)_nCT$. A stretch of trinucleotide repeats was found in the transcribed 5'-untranslated region. We located one consensus sequence on the opposite strand (TGACCTTGA) that could be recognized by the orphan nuclear receptor, steroidogenic factor-1 (SF-1, also designated AdBP4), which appears to be essential for the expression of a number of steroidogenic enzyme genes (Honda et al., *J. Biochem.* (Tokyo) **112**, 573-575 (1990); Rice et al., *Mol. Endocrinol.* **5**, 1552-1561 (1991)).

The 1.3 kb of DNA upstream from the transcription start site directed expression of the luciferase reporter gene when transfected into mouse Y-1 adrenocortical tumor cells (Example 11), but the same fragment inserted in the opposite orientation drove luciferase expression only to the same extent as the promoterless control plasmid.

In order to compare the sequence of the StAR cDNA and the sequence of an identified pseudogene, the pseudogene was also further characterized and the sequence of the pseudogene is provided in Table 7. The pseudogene sequence disclosed in the parent application contained minor sequencing errors which have been corrected in the sequence provided in Table 7.

Table 7

Nucleotide Sequence of StAR Pseudogene

GGATCTTTTTTATAGAAAACAAACTCAAGTGAGGTGGAAAATGATGAT
ATTCTTCTAATAAGAGAAAGCTCAGAAATCAGAGCTGTGAGAGTGAAA
CAGAAGGAAAGTTATGATTTAAAGACGGGTAGGCCTGATGTGATGAGA
AGCGCATTTTACTTCTGTGGTATTGTTTTCTGAAAATTTATTCACTCCA

GTTAATCATGAGAAAACAGCAGAAAAACCCAAACTGAAGGATATTCTA
CCAAATGTTTGATCAGTATAATTCAAAAGTGTCAAGCTTACAAAAAAA
TAAAGAGTGAGAACTCATAACTGGAGAACACTAGAAAATAATGCAAC
ATGGTATCATAGATTAAATACTGAAACAGAAAAAAAGGATATTAATGG
AAAAGCTGATAAAGTCTGCAAAAAGTCTGCAATTTGATTCACAGCATC
ATACGAATGTGAATTTCTAAGTTGTGATAAGTGTTTCATGGTTGCCTAC
AATGTAAACCTTAGAGAAACATGAGTAAATGGTAAGAACTCACTATAA
AATTTTGCAACTATTCTGTAAATATCCAAATAATAATAATAAAGAGGA
AATAGTAGCCAAACCAATGAAAACCAGGGAGTAATACCAAGAGTGGA
AATAAATTAAAATGGAACCAAGGGGACCAAACTACATAGACACAAAT
TAAAACTGCAACATTACCTAAATATTTCTTAAAGATATTAAGCTTTACA
TATAAAGATTATAGAAATTCATATCTACCTTGATTTTAATGACATAATG
TGTATATTAAGATTAATCTGGGTTGTTGTACATTTTCTGTATATTTCTG
AATTTGCACATTGCCAGAATGAGTAACTGGCTTGGCATTATAATTAAC
TCCTTGAGAAATTTATTTAGAGGAATAAAACAATATATTTTGCTAAGTC
ATAGAATGGACAACTCAGTTATGCTTCAGGTTATCTTAGTAGGGAGTA
TGTGGGTGAGAGGGTAACAGATATACAAATCACATCCTAGGGTTAGAC
TTACTGGGAAGATCCCATGGGATCCGAAATGGAAGTCAAAGTTTCTGT
TATCAAATTTTGGTGACTCCAAAAGGACAGGAAAGACCAGAGATAAGC
ACTAAATGAGAACAATAAATAAGCAAAAAGGTGTGTCCTACCGATTTC
AATATTCAGTGAGTCTATAAGAAGGACCTGAGCCATCGAGCCTGGCCA
AAATATTGGATTCTAATTAAAGAGTAGAGTGAGGAGGGGCACAGAGG
ACAGCCTCCAAGGGGAGGCCGCACTGCAAGCATCCCTGGAGTGGCGA
AGGTATGCACTGGATGGATGGCAGCAGGCGCTGCACGGGGGAGCTGA
GCACTGCCAGGAAGAATCCAGTGAGTGATGGCGTTTATATCTCCTGAT
GATGATTCACAGCCTTCAGTGGGGGACATTTAATACGTGGAACACCGG
GTCCAGGCTGCAGCTGCGGGACTCAGAGGCAAAGCTTGAGTGGCTCAG
GAAGGACGAAGAACCACCCTTGAAAGAAGAGGCAGCCTCACCGGCGT
TGGCGGCCCCACCACTGCCACATCTGCCAGGAAAGAATGCTGCTAGCG
ACATTCAAACTGTGCTCCAGGAGCTCCTACAGACACATGCGCAACATG
AAGGGGCTGAGGCAACAGGCTGTGAGGGGGCATCGGGCAGGAGCTTA

ACCGGAGGGCCCTGGGGGCCCCACCCCAAGCGCTTGGATTAACCAGGT
TCCGCCGCGAGCTCTCTTGTTCTCTCTGGAAGAGACTCTCTACCCGGGT
GCGGTGGCTCACGCCTGTAATACTAGCACGTTGGGCCGAGGCGGGCAG
ATCATGAGGTTAGGAGTTCGAGAGCAGNCCGACCCACATGGTGAAACC
CCATCTCTACTAAAAATACAAAAATTAGCTGGGAGTGGTGGTGCGGGC
CTGTAATCCCAACTACTCAGGAGGCTGAGGCAGGAGAATCGCTTGAAC
TCGGGGACGGGGGGGCGGGCGGGGAAAGACTCTCTACAGTGACCAGG
AGCTGACCTATCTCCAGCAGTGGGGAGGAGGCCATGCAGAAGGCCTTG
GGCATCCTTAGCCCTCGCCAACTACGAGGGCTGGAAGAAGGAGAGCCA
CCAGGACAATGGGGATAAGTATAGTAAAGTGGTTCCAGATGGGGCAAG
GTGTTCCGGCTGGAAGTCGTGGTGGACCAGCCCATGGAGAGGCTCTAC
AAAGAGCTCGTGGAGTGCATGGAGGCAATGGGGGAGTGCAACTCCAA
TATCAAGGCGATCAAGGTCTTGCAGAAGATGATCAGAAAAGATACATT
CATTGCCCATGAGCTGGCTGCAGAGGCAGCAGGAAACCTAGTGGGGCC
TTGTGACTCTGTGAGCATGTGCTGTGCCAAGCGTCAAGGCTCCACCTG
TGTGCTGGCTGGCATGGCCACAGACTTCGGGAACATGCCCGAGCAGAA
GGGTGTCATCAGGGGGAGCATGGTCCCACTTGCATGGTGCTTCACCTG
GTGACTGGAAGTCCCTCCAAGACCAAACTTACATGACTGCTCAGCATC
GACCTCAAGGGGTGGCTTCCCAAGAGCATCATCAACCAGGTCCTGTCC
CAGACCCAGGTGGATTTGGCCAACCACCTGCACAAGCGCCTGGAGTCC
CACCCTGCCTCTGAAGCCAGGTGTTGAAGGCCAGCCTGCTGTTCCCAA
GTGTGTCCAGCTGCACTGCTACACACGCTTATCAGGAGAATCCTTGCT
GGAAGCCTGCAAGCTTAAAATCTCCATCTGGCGACAGAGGAATAGGTG
GGGTTAGTGTATAGAGTATGATACTAGGATTCAGACTGGTAAAAGTTT
TTAGTACCAAGAAAACAAGGATGAGGCTCTTTGATTAAAAGGTAACTT
CATTCACTGACTAGCTATGACATGAAGGTTGAGGATCCTAAAATAATT
GTAAAACTTTTTTTCTGGGCCTTTATGTGCCCACCTAAAACCATCTTTA
AAATGCTAGTGGCTGATATGTGTGGGGGGATGCTAGTCACAGGGCCTG
AGGAGTCTTGCTTTATGGGCTGGAGAACCCCATGCCCTGAAGGCAGAG
CATGTGCACAAAGCAGAATCTTAGAGGGTCTCCTGCAGCCCTCCACTC
CTCCAAGTCGCTGCATGGCAACACCAGATAACAAGCAGCACCCCACAG

TGGGTACCTTCCAGAAATATAGTCCAAGCTTTCTCTATGGAAAAAGAC
AAAACTAATTAGTAAATAGGTTTCCCTATTGAGTCCATAGGCACCAGT
CAGAAAAAAGAATCATAATTCACACACACAAACACACACACACACAC
ACACACAAAACAAGGACCTGAGTTCAGAAAATGAAGCCTGTAATCACA
CACTAAAATGAAAACAATAAATCATGTGTAAACAGTTAATAAATGAAT
AAAATGTATTGCTTCTATAGCCTTGTGATATGGTTTGGCTGTGTCTGCA
CCCAAATCTCATCTT

Thus, the present invention provides an isolated DNA molecule, in which the molecule contains (1) a first sequence consisting of hStAR cDNA, hStAR genomic DNA, hSTAR promoter or hStAR pseudogene as set forth in Figure 1, Table 6, or Table 7; (2) a second sequence, wherein the second sequence is a subsequence of the first sequence at least 10 nucleotides in length; (3) a third sequence in which at least one nucleotide of the first or second sequence is replaced by a different nucleotide; (4) a fourth sequence in which at least one nucleotide is deleted from or inserted into said first or second sequence; or (5) a fifth sequence complementary to any of the first, second, or third sequences; with the provisos that (1) said molecule can be an RNA molecule in which U replaces T in any of said sequences (1) - (5), (2) the third sequence is at least 95% identical to the first or second sequence, (3) the second sequence is not present in mouse StAR cDNA, and (4) said fourth sequence contains no more than 20 inserted nucleotides and no more than 200 deleted nucleotides. Any of these sequences can be used in the identification of the presence (or absence) of a mutation in the StAR gene of a human and thus can be used in the genetic counseling of individuals, for example those with a family history of congenital lipoid adrenal hyperplasia (although the general population can be screened as well). In particular, it should be noted that the invention is not limited to use or identification of the specific mutations that have already been identified. Any mutation in the StAR gene away from the normal gene sequence identified here is an indication of a potentially fatal genetic flaw, even so-called "silent" mutations that do not encode a different amino

acid at the location of the mutation are potential disease mutations, since such mutations can introduce into (or remove from) the gene an untranslated genetic signal that interferes with the transcription or translation of the gene. Since one of the utilities based on the gene sequences identified here is in genetic counseling of families with a history of lipoid CAH, advice can be given to a patient concerning the potential for transmission of lipoid CAH if any mutation of the StAR gene is present. While an offspring with the mutation in question may or may not have symptoms of lipoid CAH, patient care and monitoring can be selected that will be appropriate for the potential presence of the disease; such additional care and/or monitoring can be eliminated (along with the concurrent costs) if there are no differences from the normal gene sequence. As additional information (if any) becomes available (e.g., that a given silent mutation or conservative replacement mutation does or does not result in lipoid CAH), the advice given for a particular mutation may change. However, the change in advice given does not alter the initial determination of the presence or absence of mutations in the StAR gene that this invention has for the first time indicated to be a sufficient cause of lipoid CAH.

Molecules containing the full-length StAR cDNA sequence are useful as sources of subsequences (discussed below) or as starting materials for the preparation of the StAR molecule itself. A "subsequence" is a group of consecutive nucleotides from one of the indicated full-length sequences. Such subsequences can be prepared by chemical synthesis from starting nucleotides (as in an automated gene synthesizer) or by biochemical manipulation of the full-length sequences (e.g., using restriction endonucleases to prepare fragments, optionally followed by (1) cleavage of terminal nucleotides with exonucleases and/or (2) size sorting and/or affinity capture to select the desired sequence). Any subsequence of the StAR cDNA sequence of sufficient length to be unique under the conditions being used is useful as one of the two primers used in a polymerase chain reaction (PCR) amplification of all or part of the genomic StAR gene as part of a method of identifying the presence or absence of a given StAR gene mutation, such as those described in this specification; the second primer is simply selected from the

opposite strand sequence so that the mutation or other sequence to be amplified lies between the two primers. Another preferred subsequence is one that contains a mutation from the normal sequences described herein, as such sequences can be used in allele-specific hybridization techniques to detect the presence of specific mutants. Preferred subsequences also include those that can distinguish between the normal StAR gene and the pseudeogene (i.e., that are not found in both the normal StAR gene DNA of Table 6 or the StAR pseudogene DNA of Table 7 or that span the alternative splice region shown in Figure 5).

The length of a subsequence necessary to uniquely hybridize with the desired target sequence will vary with the particular method being used and is within the ordinary skill of those who carry out routine identification of genetic material. Typical primers are at least 10, preferably at least 14, more preferably at least 17, even more preferably at least 20 nucleotides in length and typically no more than 200, preferably no more than 100, more preferably no more than 70, even more preferably no more than 50 nucleotides in length. The most preferred subsequences are found in at least one of the human StAR sequences set forth in Table 6 and 7 but are not found in mouse StAR DNA.

In addition to those molecules that contain sequences and subsequences identical to the those of the StAR gene, molecules containing mutated sequences are also useful, as they can be used as specific probes for mutations. For example, several mutations of amino-acid-encoding codons into stop codons (i.e. nonsense mutations) are identified in the following examples and elsewhere in the specification; e.g., $Arg^{193} \rightarrow Stop$ and $Gln^{258} \rightarrow Stop$ mutations. (Here and elsewhere in this specification "codon" refers to a nucleic acid triplet in the reading frame of the gene, unless otherwise clear from the context.) Thus, a preferred class of mutant-sequence molecules are those that contain a replacement (or more than one replacement) of a nucleotide that converts a codon to a stop codon at a location other than the 3' terminus of the coding sequence, so that a truncated, non-functional StAR polypeptide molecule is encoded. The mutated codon is located preferably at least 5, more preferably 10, even more preferably 20, still more preferably 30 codons distant from the 3' terminus of the normal coding sequence

so that sufficient deletion would occur in the targent to produce a non-functional product. Another preferred class of mutant-sequence molecules contains a deletion of at least one nucleotide, preferably 5 nucleotides, more preferably over a hundred nucleotides, and most preferably up to 185 nucleotides or more, but not exceeding 200 nucleotides. Another preferred class of mutant-sequence molecules contains an insertion of no more than 20 nucleotides, preferably less than 5 nucleotides, and most preferably one nucleotide. Other preferred classes of mutant-sequence molecules are those known to produce non-functional StAR molecules, such as those resulting in non-conservative amino acid replacement, and those that alter translation or transcription signal sequences present in the gene or that introduce improper translation or transcription signal sequences.

It will be recognized that the discussion immediately above refers to sequences and subsequences in the sense strand of genomic DNA. Such sequences can be used to detect the presence of the anti-sense strand of genomic DNA as a result of their complementary nature. However, it is also possible to use a sequence complementary to any of those discussed above, since they will be complementary to and detect the sense strand.

Molecules of the invention will contain a sequence that is different from the mouse genomic StAR gene sequence (in the region from the initiation codon to the stop codon for the StAR gene product) and at least 95% identical to the human StAR cDNA or genomic sequence. By 95% identical is meant that the sequence in question contains no more than 5% different nucleotides from the sequence to which it is being compared, counting each insertion, deletion, or substitution of a nucleotide as a single difference. It will be apparent that a sequence less than 20 nucleotides in length will have to be identical to the standard sequence if it is to be greater than 95% identical.

Identity and relative identity can readily be understood by reference to the following examples. For example, if the hypothetical sequence

abcdabcdabcdabcdabcdabcdabcdabcdabcd,

which is 40 "nucleotides" in length, is considered to be the standard against which a measurement is being made, each of the following hypothetical nucleotide

sequences is 95% identical to the standard sequence (i.e., each has two single-nucleotide differences from the standard 40-nucleotide sequence):

abcdabcdabcdabcdabcdabcdabcdabcdabcdab
[two deletions at 3' terminus];

abcabcdabcdabcdabcabcdabcdabcdabcdabcd
[two random-location deletions];

ababcdabcdabcdabcdabcdabcdabcdabcdabcdabcd
[two insertion at 5' terminus];

abcdabcdabcdabdabcdabcdabcdabcdaabcdabcd
[one random insertion and one random deletion];

abcdabcdbbcdabcdabcdabcdabcdabcdbbcdabcd
[replacement of two "a" nucleotides by "b" nucleotides]; and

abcdabcbabcdabcdabcdabcdabcadabcdabcdabcd
[one replacement and one insertion].

It will be apparent that many similar examples could be given, particularly with molecules of the invention, which are often of larger size than these examples. However, these examples should suffice to teach a person of ordinary skill the meaning of "% different" as used herein. It will also be readily recognized that the sequences to be compared will be aligned for maximum identity before differences are calculated; while computer programs (such as the FASTA program, described in Pearson, W.R., and Lipman, D.J., *Proc. Natl. Acad. Sci. USA*, 85 2444-2448 (1988)) can be used, the high degree of required homology means that visual sequence comparisons will readily find the maximum homology alignment.

In a preferred embodiment, gap widths of up to 200 base pairs are allowed in the alignment program.

The specific sequences indicated above to be derived from or otherwise related to a StAR gene can be the entire sequence of a polynucleotide or can be part of a larger sequence. For example, sandwich hybridization assays that utilize long polynucleotide sequences containing subsequences that hybridize with different molecules (such as target genomic sequences or sequences present in a second polynucleotide that acts as an anchor to a solid surface) are well known. See, for example, U.S. patent Nos. 5,288,609 and 5,124,246.

The word "isolated," when used to refer to a polynucleotide molecule characterized by the sequences set forth in this specification, means separated from at least some of the genomic DNA normally associated with the StAR gene and preferably separated from all human cellular materials other than polynucleotides. Gene libraries that may have contained a vector containing an unidentified segment of genomic DNA including the StAR gene are not "isolated," as the StAR gene was not known to be present and/or was not separated from vectors containing other human DNA. In most cases, an isolated molecule of the invention will have a length of less than 50 kb, preferably less than 30 kb, more preferably less than 20 kb. Minimum lengths have been previously discussed.

Generally, the compositions of the invention will be used in a method of detecting the presence of a genetic defect that causes or may cause congenital lipoid adrenal hyperplasia in a human or that can or may transmit congenital lipoid adrenal hyperplasia to an offspring of the human, in which the compositions are used to identify a mutation of a StAR gene of the human. Initially, genetic counselors and others will be looking simply for differences from the StAR gene sequence now identified as being normal and not associated with disease, since any deviation from this sequence has the potential of causing disease, which is a sufficient basis for genetic counseling, particularly if the different (but still unconfirmed) gene is found in a person with a family history of congenital lipoid adrenal hyperplasia. As specific mutations are identified as being positively correlated with congenital lipoid adrenal hyperplasia (or its absence), genetic

counselors will in some cases focus on identifying one or more specific mutations of the StAR gene that changes the sequence of a protein product of the StAR gene or that results in the StAR gene not being transcribed or translated. However, simple identification of the presence or absence of any mutation in the StAR gene of a patient will continue to be a viable part of genetic analysis and counseling.

Q258X and R182L account for 70-80% of affected alleles in the Japanese and Palestinian populations, respectively, providing the opportunity for efficient genetic screening. Thus we have devised PCR-based tactics to facilitate diagnosis of these and other mutations (Figure 8). In each case genomic DNA is amplified with a pair of primers that encompasses the suspected mutation, and the PCR product is then cut with a restriction endonuclease whose recognition sequence is created or destroyed by the mutation. Table 4 lists the oligonucleotide pairs, restriction endonuclease and cleavage pattern for the Q258X and R182L mutations, as well as for the six other mutations that could be diagnosed similarly. The sequences of all of the oligonucleotides are provided in Table 8.

Table 8

Oligonucleotides used for PCR amplification of genomic DNA

Primer#	Sequence (5' → 3')	Purpose	PCR Program	Size of Product
Ex1S-5 Ex1S-L Ex1AS	GCAGCAGCAGCGGCGGCAGCAG TAACACAGGTTTCTGAGCCTCAAT ATCAGAATTGGGTGGCCTGAGCCTC	Sequence Exon 1	3	621
Ex2S Ex2AS	GTCCCTGCTAGAATACTGTGTT AAAGCCACATGCACCACATCA	Sequence Exon 2	2	342
Ex3S Ex3AS	CAATGAGCAGACCCAGAGCT GACTGCTGCATGAGACAGGA	Sequence Exon 3	2	308
Ex4S Ex4AS	TGCTGGGATTATAGGCGTGAAC GCTAGGGGTCCTCTCTTTGATACAG	Sequence Exon 4	2	300
S3 AS1	GTGAGCAAAGTCCAGGTGCG ATGAGCGTGTGTACCAGTGCAG	Sequence Exon 5-7	1	2081
Ex5S Ex5AS	TGCTGTATCAAAGAGAGGAC AGCCTGCTGCCCGTATTTAC	RFLP Exon 5	4	295
S4 AS4	CCTGGCAGCCTGTTTGTGATAG CCTCATGTCATAGCTAATCAGTG	RFLP Exon 7	1	439
B2	GACCACAAGATGAGCACATTC	RFLP Exon 7	4	205
S3 AS2	GTGAGCAAAGTCCAGGTGCG TGTGGCCATGCCAGCCAGCA	Sequence Intron4/ Exon5 Junction		432

The actual technique used to identify the StAR gene or a StAR gene mutant is not itself part of the practice of the invention. Any of the many techniques that can be used to identify gene mutations, whether now known or later developed, can be used, such as hybridization with specific probes, which includes the technique known as allele-specific oligonucleotide hybridization (either without amplification or after amplification of the region being detected, such as by PCR), restriction fragment length polymorphism (RFLP) analysis, or random amplified polymorphic DNA (RAPD) analysis. Other analysis techniques include enzymatic mismatch

scanning and transcription/translation analysis. All of these techniques are described in a number of patents and other publications; see, for example, for RFLPs, D. Botstein et al., in the *American Journal of Human Genetics* **32**, 314-321 (1980), and for RAPDs, J.G.K. Williams et al., in *Nucleic Acids Research* **18**, 6531-6535 (1990).

Depending on the patient being tested, different identification techniques can be selected to achieve particularly advantageous results. For example, for a group of patients belonging to a particular racial or ethnic group known to be associated with a particular mutation of the StAR gene, allele-specific oligonucleotide (ASO) hybridization is a preferred technique. For screening of large, mixed-origin populations, single-strand conformation polymorphism is preferred. For an individual, total sequencing of genetic and/or cDNA and comparison with standard sequences, such as those shown herein, are preferred.

In many identification techniques, some amplification of the host genomic DNA (or of messenger RNA) will take place to provide for greater sensitivity of analysis. In such cases it is not necessary to amplify the entire StAR gene, merely the part of the gene or the specific location within the gene that is being detected. Thus, the method of the invention generally comprises amplification (such as via PCR) of at least a segment of the StAR gene, with the segment being selected for the particular analysis being conducted by the diagnostician.

Since lipoid CAH is an autosomal recessive genetic disease, the method of the invention in some cases will classify the patient as homozygous for the normal StAR gene or for the mutated StAR gene or heterozygous for the normal StAR gene and the mutated StAR gene, since this information is informative for genetic counseling.

The patient on who diagnosis is being carried out can be an adult, as is usually the case for genetic counseling, or a newborn, or prenatal diagnosis can be carried out on a fetus. Blood samples are usually used for genetic analysis of adults or newborns (e.g., screening of dried blood on filter paper), while samples for prenatal diagnosis are usually obtained by amniocentesis or chorionic villus biopsy.

The full-length normal StAR genes from humans, as well as shorter genes that produce functional StAR proteins, can be used to correct congenital lipoid adrenal hyperplasia in a human patient by supplying to the human an effective amount of a gene product of a human StAR gene, either by gene therapy or by in vitro production of the StAR protein followed by administration of the protein. Since lipoid CAH is recessive and is thus treatable by supplementary supply of StAR, such treatment is readily accessible. Likewise, treatment or prevention of hypercholesterolemia can be achieved by supplying to the human liver an effective amount of the gène product of the human StAR gene, either by gene therapy or by administration of a StAR protein produced *in vitro*.

It should be recognized that the various techniques for administering genetic materials or gene products are well known and are not themselves part of the invention. The invention merely involves supplying the genetic materials or proteins of the invention in place of the genetic materials or proteins previously administered. For example, techniques for transforming cells to produce gene products are described in U.S. Patent No. 5,283,185 entitled "Method for Delivering Nucleic Acid into Cells," as well as in numerous scientific articles, such as Felgner et al., "Lipofection: A Highly Efficient, Lipid-Mediated DNA-Transfection Procedure," *Proc. Natl. Acad. Sci. U.S.A.*, 84 7413-7417 (1987); techniques for in vivo protein production are described in, for example, Mueller et al., "Laboratory Methods - Efficient Transfection and Expression of Heterologous Genes in PC12 Cells," *DNA and Cell Biol.*, 9(3), 221-229 (1990). Administration of proteins to overcome a deficiency disease is so well known (e.g., administration of insulin to correct for high blood sugar in diabetes) that further discussion of this technique is not necessary. Some modification of existing techniques may be required for particular applications, but those modifications are within the skill level of the ordinary practitioner using existing knowledge and the guidance provided in this specification.

The StAR promoter can be used to drive the expression of the StAR gene, a StAR mutant or a heterologous gene in a mammalian cell, in a transgenic mammal or for expression of a heterologous gene in gene therapy. Tissue specific

gene expression can be achieved to the extent that the heterologous gene would not be expressed in many of those cell types that do not normally produce StAR mRNA, such as, for example, placenta, liver, and choriocarcinoma cells and can be expressed in cells that normally produce StAR mRNA, such as, for example, adrenal and gonadal cells. A preferred promoter fragment includes the sequence A^{-1300} GCTT to GGTCC^{-1} of Figure 7 and smaller fragments can be used, but these are less preferred.

The invention now being generally described, the same will be better understood by reference to the following detailed examples, which are provided for purposes of illustration only and are not to be considered limiting of the invention.

EXAMPLES

Example 1: Isolation of human StAR cDNA clones and DNA sequence analysis

A human adrenal cortex cDNA library in lambda gt22A, prepared from poly (A)+ RNA isolated from the adrenal cortex of an 18 year old male, was provided by Drs. Andre Lacroix, Alain Belanger, and Yves Tremblay, University of Laval; Quebec, Canada. The library was screened with a partial-length mouse StAR cDNA (Clark et al., 1994). More than 50 positive clones were detected in the screening of 600,000 plaques. Two plaque-purified phage clones were selected for sequence analysis. Each contained an insert of approximately 1.6 kb. Both inserts were subcloned into pSPORT (GIBO-BRL, Bethesda, MD) and sequenced utilizing an automated DNA sequencer (Applied Biosystems, Inc.) employing Taq dideoxy sequencing reagents. Ambiguities were corrected by manual sequencing.

The two human StAR cDNAs that were characterized by DNA sequence analysis had identical 126 nt 5'-untranslated regions. Both clones contained an 855 nt open reading frame encoding a 285 amino acid protein. The 1.6 kb cDNA whose nucleotide sequence is shown in Figure 1 had a 623 nt 3'-untranslated sequence that ended in a poly (A)+ tail preceded 23 nt upstream by an AATAAA sequence.

The deduced human StAR amino acid sequence is 84% identical to that of mouse StAR (Clark et al., 1994) (Figure 1). It contains a 25 amino acid N-terminal sequence that is comprised of basic and hydrophobic amino acids that are characteristic of mitochondrial targeting sequences. Seven consensus sites for phosphorylation by cAMP-dependent protein kinase and three protein kinase C phosphorylation sites are present in the sequence of the mature protein. Expression of StAR in engineered COS-1 cells increases steroidogenesis.

Example 2: Expression of StAR cDNA in COS-1 cells

To examine the functional activity of the human StAR protein, we utilized methods that we previously employed to explore the function of sterol carrier protein 2 in steroidogenesis (Yamamoto, R., Kallen, C.B., Babalola, G.O., Rennert, H., Billheimer, J.T., Strauss, III J.F. (1991) Proc. Natl. Acad. Sci. USA 88: 463-467). Briefly, COS-1 cells were transfected with various expression vectors with Lipofectamine (GIBCO-BRL) using 10 μl/dish. The vectors included pSPORT without cDNA insert, pSPORT with the 1.6 kb StAR cDNA (pStAR), and expression vectors for bovine P450scc (pCDP450scc) and adrenodoxin (pCDADX), provided by Dr. Michael Waterman, Vanderbilt University (Nashville, TN). Forty-eight hours after transfection, medium was collected for radioimmunoassay of pregnenolone as previously described (Yamamoto, R., Kallen, C.B., Babalola, G.O., Rennert, H., Billheimer, J.T., Strauss, III J.F. (1991) Proc. Natl. Acad. Sci. USA 88: 463-467). In one experiment, the hydroxysterol, 20α-hydroxycholesterol, was added (5 μg/ml) to the incubation medium. This hydroxysterol is a more soluble pregnenolone precursor and an intermediate in the cholesterol side-chain cleavage reaction. Hydroxysterols, like 20α-hydroxycholesterol, by-pass the regulated translocation mechanism of cholesterol movement and, therefore, generally provide an index of maximal cholesterol side-chain cleavage activity (Toaff, M.E., Scleyer H., Strauss, J.F., III (1982) Endocrinology 1785-1790). Preliminary studies established that the transfected COS cells secreted about 10-fold more pregnenolone than progesterone and that the measured progesterone levels changed in parallel with the

pregnenolone. Consequently, we monitored pregnenolone secretion as our index of steroidogenic response.

COS-1 cells did not secrete pregnenolone when transfected with the pSPORT vector lacking a cDNA insert or the pSPORT vector harboring the StAR cDNA (Table 9). However, co-transfection of the cells with plasmids directing expression of bovine P450scc and adrenodoxin endowed the cells with steroidogenic activity. Triple transfection of the COS-1 cells with P450scc, adrenodoxin and StAR expression plasmids consistently increased steroid secretion 4-to-20-fold over cells transfected with P450scc, adrenodoxin and the control pSPORT plasmid. Incubation of cells transfected with pP450scc, pADX and pSPORT with 20α-hydroxycholesterol, a relatively soluble intermediate of the cholesterol side-chain cleavage reaction, stimulated pregnenolone secretion to the same extent as pStAR but did not augment the pStAR response in COS cells co-transfected with P450scc and adrenodoxin plasmids. In the absence of P450scc and adrenodoxin expression, there was no detectable pregnenolone synthesis in the presence of 20α-hydroxycholesterol. These findings document that the pSPORT plasmid "control" did not interfere with expression of the steroidogenic enzymes. The fact that an exogenous hydroxycholesterol did not augment steroid production stimulated by StAR also suggests that StAR promotes nearly maximal steroidogenic activity in the transfected COS cells.

The more than 4-fold increase in steroidogenesis promoted by expression of StAR in the COS cell system is substantially greater than the 2-fold increase we observed when COS cells were transfected with sterol carrier protein 2 expression plasmids as the vehicle for enhancement of steroidogenesis (Yamamoto, R., Kallen, C.B., Babalola, G.O., Rennert, H., Billheimer, J.T., Strauss, III J.F. (1991) Proc. Natl. Acad. Sci. USA 88: 463-467). While these observations are consonant with the idea that StAR facilitates steroidogenesis, these studies do not define the exact mechanism of StAR action.

Table 9

Stimulation of steroidogenesis by StAR in COS-1 cells transfected with cholesterol side-chain cleavage enzyme and adrenodoxin.

Treatment	Pregnenolone secretion (ng/dish)			
	Expt. 1	Expt. 2	Expt. 3	Expt. 4
Mock transfection	<5			
pSPORT	<5			
pStAR	<5			
pStAR+20α-OH-C				<5
pCDP450scc+pCDADX+pSPORT	26 ± 6	14 ± 1.0	10 ± 0.01	20 ± 0.5
pCDP450scc+pCDADX+pSPORT+20α-OH-C				157 ± 10
pCDP450scc+pCDADX+pStAR	545 ± 50	78 ± 4	41 ± 2.0	175 ± 10
pCDP450scc+pCDADX+pStAR+20α-OH-C				137 ± 8

COS-1 cells were transfected with the indicated plasmids (2 μg plasmid/35 mm dish) with Lipofectamine. The media were collected after 48 h and assayed for pregnenolone by radioimmunoassay. 20α-hydroxycholesterol (20α-OH-C; 5 μg/ml) was added to some cultures. The results of 4 separate experiments are presented. Values are means ±S.E., N=3-4 replicates per experiment.

Example 3: Identification of StAR Mutants

Genomic DNA prepared from leukocytes was amplified using various primers (Table 8) according to one of four empirically determined PCR programs, all of which were initiated by denaturation at 95°C for 2 min. Program 1: 34 cycles of 94°C for 50s, 64°C for 30s, 72°C for 90s; Program 2: 94°C for 45s, 57°C for 45s, 72°C for 60s; Program 3: 94°C for 50s, 60°C for 45s, 72°C for 90s; Program 4: 29 cycles of 94°C for 60s, 55°C for 60s, 72°C for 120s. All programs used Taq polymerase from Perkin-Elmer-Cetus and were terminated by a final extension of 72°C for 15 min. PCR products were separated by agarose gel electrophoresis and purified with either Gene Clean (BIO 101) or Qiagen resin (Qiagen). Exons 1-4 were amplified with the primer pairs shown in Table 8 and sequenced directly on an Applied Biosystems automated sequencer, as described

(Ohba, et al. 1994). Exon 1 was amplified for automated sequencing using Ex1S-S (short) and Ex1AS and for manual sequencing using Ex1S-L (long) and Ex1AS. The 2.1 kb PCR fragment containing exons 5-7 was amplified with primers S3 and AS1 and cloned into pCRII (Invitrogen) isolated by Wizard minipreps (Promega) and sequenced using ^{35}S[dCTP] by dideoxy chain termination with Sequenase 2.0 (USB). Ambiguities from automated sequencing were resolved by manual sequencing; all manual sequencing was performed on both strands. Some of the mutations were confirmed by restriction endonuclease digestion of PCR amplified DNA as described in Table 3.

Example 4: Activity of the StAR Mutants

To determine if the identified mutations were indeed the cause of the patients' lipoid CAH and to begin to elucidate the structure/function requirements of the StAR protein, we tested each identified mutant *in vitro*. The various StAR mutations were re-created by site-directed mutagenesis, cloned into StAR expression vectors in either pMT2 or pSPORT, and transfected into non-steroidogenic COS-1 cells. Normal human StAR cDNA was cloned into the *Eco*RI/*Hind*III site of expression vector pCMV4 (Andersson, et al.). Identified mutations were re-created by overlapping PCR using the sense and antisense primers FL-S and FL-AS (which correspond to the 5' and 3' ends of Full-Length StAR cDNA and contain *Eco* RI and *Hind*III sites respectively) and one of the paris of complementary primers containing the sequences to be altered (Table 10). Two separate PCR reactions were done using PCR program 4 and PFU polymerase (Stratagene) for the construction of each mutation: the first used FL-S and the antisense mutagenic oligonucleotide; the second used FL-AS and the sense mutagenic oligonucleotide. These PCR products were purified by agarose gel electrophoresis followed by Geneclean or Quiagen columns. After a 100-fold dilution, a 1μl aliquot of each pair of reactions was combined and used as template for PCR (Program 4) with the FL-S and FL-AS primers. The final PCR product was cut with *Eco* RI and *Hind*III-cleaved pCMV4 and sequenced to verify the accuracy of the construction. The ΔR272 and 947/InsA/948 mutations, which lie

close to the 3' end, were constructed in single segments using oligonucleotides FL-S and either ΔR272-AS or 947/InsA-AS for the first PCR reaction, and FL-S and the 3' Bridge oligonucleotides for each of the second reactions, all using Program 4.

Table 10

Oligonucleotides used for Mutagenesis

Name	Sequence 5' → 3'	Purpose
FL-S	TGCTGAATTCATGCTGCTAGCGACATTC	Full length StAR cDNA
FL-AS	AGCTAAGCTTTGGTCTTCAACACCTGGCT	"
E169G-S	TACTCACGGGCTGGCTGCCGAGGCA	Create E169G
E169G-AS	CAGCCAGCCCGTGAGTAATGAATGTATC	"
R182L-S	TGGGGCCCCTTGACTTTGTGAGCGT	Create R182L
R182L-AS	CACAAAGTCAAGGGGCCCCCACCAGGTT	"
ΔT593-S	GAGATCAAGG_CCTGCAGAAGATCGG	Create ΔT593
ΔT593-AS	TTCTGCAGG_CCTTGATCTCCTTGACA	"
ΔC650-S	CCGAGGCAG_AGGAAACCTGGTGGGGCC	Create ΔC650
ΔC650-AS	GGTTTCCT_CTGCCTCGGCAGCCAGCTC	"
ΔR272-AS	CAGGCGCTT_ _ _CAGGTGGTTGGCAAAATC	Create ΔR272
947/InsA-AS	CTCCAGGTCGCTTGCGCAGGTGGTT	Create 947/InsA/948
Bridge	ACACCTGGCTTCAGAGGCAGGGTGGGACTCCAG	

The mutated bases are indicated by underlining (mis-sense) mutations or blanks (base deletions).

To assay the effect of normal or mutant StAR proteins on the conversion of cholesterol to pregnenolone, the cells were co-transfected with a vector expressing the three components of the cholesterol side-chain cleavage system as a single monomolecular fusion protein (H_2N-P450scc-AdRed-ADx-COOH) termed F2 (Harikrishna, et al. 1993). This optimizes P450scc activity and

obviates any variations in P450scc activity due to variation in the molar ratio of P450scc to its electron transfer proteins, especially adrenodoxin (Harikrishna, et al.; Zuber, et al. *Proc. Natl. Acad. Sci. USA* **85**, 699-703 (1988)). Incubation of cells with a soluble cholesterol analogue such as 20α-hydroxycholesterol or 22R-hydroxycholesterol bypasses the mitochondrial steroidogenic capacity (Toaff, et al. *Endocrinology* **111**, 1785-1790 (1982)). Thus the ratio of steroidogenic capacity with LDL cholesterol as substrate versus 22R-hydroxycholesterol as substrate provides an index of the efficiency of mitochondrial cholesterol transport. As seen in Table 11, the wild-type StAR caused a 10-fold increase in steroidogenesis, consistent with our previous findings (Lin, et al. 1995), and none of the mutants had activity greater than the vector control. Thus each of the identified mutations yielded a wholly inactive StAR protein.

Table 8

Activity of StAR Mutants

Plasmid	Activity[a]	(% of wild type)
Control Vector	0.059	14 ± 2
Wild type StAR	0.433	100
Amino Acid Changes:		
E169G	.043	11 ± 2
R182L	.033	8 ± 2
ΔR272	.043	10 ± 1
E169K	.062	14 ± 2
A218V	.088	20 ± 4
L275P	.103	24 ± 5
Frameshifts:		
ΔT593	.035	9 ± 3
ΔC650	.042	10 ± 2
947/InsA/948	.041	10 ± 2
Stop Codons:		
Q258X	.070	16 ± 4

[a] Activity is calculated from ratio of steroidogenic activity with LDL cholesterol substrate to 22R-hydroxycholesterol substrate, measured by immunoassay of secreted pregnenolone. The value with the normal StAR control is set at 100% values are means ± SEM of three separate experiments, each performed in triplicate.

Example 5: Expression of StAR mRNA

Northern blots containing 2 μg of poly (A)+ RNA from various human tissues were purchased from Clontech Laboratories (Palo Alto, CA) and probed with the 1.6 kb StAR cDNA and a β-actin cDNA according to the supplier's protocol.

StAR mRNA was detected in human ovary, testis and kidney. The most abundant transcript was 1.6 kb and less abundant mRNAs of 4.4 and 7.5 kb were observed in ovary and testis (Figure 9). The ovarian sample, prepared from a pool of five ovaries obtained from women of reproductive age, contained the most StAR mRNA followed by the testis and then the kidney. In the Northern blots shown in Figure 9, probed simultaneously with the same preparation of ^{32}P-labeled cDNAs, the blot containing the ovary and testis was exposed for 6 h for expression of StAR whereas the blot containing the kidney sample was exposed for 24 h for StAR. Longer exposures of both blots failed to reveal StAR mRNA in placenta, pancreas, skeletal muscle, liver, lung, brain, heart, peripheral blood leukocytes, colon, small intestine, prostate, thymus and spleen. However, β-actin mRNA was readily detected in all of these tissues on the same blots. StAR expression in human adrenal cortex is inferred from the fact that multiple StAR phage clones were detected in the library used to isolate the human StAR cDNA.

These observations suggest that StAR expression is restricted to organs that carry out mitochondrial sterol hydroxylation reactions that are under acute regulation by tropic hormones that act via the intermediacy of cAMP. This is true for the adrenals and gonads, which respond to their respective pituitary tropic hormones, ACTH and LH, with enhanced cholesterol side-chain cleavage, and to the kidney, which increases 1α-hydroxylation of vitamin D in response to PTH. It is notable that another steroidogenic organ, the placenta, does not appear to express StAR. However, placental progesterone does not seem to be under acute regulation by cAMP. The reported stimulatory effect of agents that raise placental trophoblast cAMP levels or cAMP analogs is most likely related to increased expression of genes encoding steroidogenic enzymes, a process that takes hours or days (Golos, T.G., Miller, W.L., Strauss, III, J.F. (1987) J. Clin. Invest. 80: 896-899). The brain, which is also a site of steroidogenesis (Patterson, D., Jones, C., Hart, I., Bleskan, J., Berger, R., Geyer, D., Eisenberg, S.P., Smith, M.F., Jr., Arend, W.P. (1993) Genomics 15: 173-176), did not appear to express StAR either. The absence of StAR expression in the placenta and brain suggests that steroid hormone synthesis in these organs is regulated by other mechanisms, a

suggestion that has been previously made by Lieberman and colleagues (Lieberman, S., Prasad, V.V.K. (1990) Endocr. Rev. 11: 469-493).

Total RNA was also isolated from cultures of human granulosa cells obtained from women undergoing in vitro fertilization/embryo transfer, or from purified human cytotrophoblast cells. The human granulosa cells were cultured for 4 days and then treated with 1.5 mM 8-bromo-cAMP for 24 h. The cytotrophoblast cells were cultured for 24 h in the absence or presence of 1.5 mM 8-bromo-cAMP. Detailed protocols for the preparation, culture and isolation of total RNA from the granulosa cells and trophoblast cells have been described previously (Golos, T.G., Miller, W.L., Strauss, III, J.F. (1987) J. Clin. Invest. 80: 896-899; Ringler, G.E., Kao, L.-C., Miller, W.L., Strauss, III, J.F. (1989) Mol. Cell. Endocrinol. 61: 13-21). Northern blots were probed with the StAR cDNA and a cDNA encoding human 28 S rRNA.

Culture of human granulosa cells in the presence of 1.5 mM 8-bromo--cAMP for 24 h increased StAR mRNA 3-to 7-fold relative to 28 S rRNA (Figure 10). In contrast, StAR mRNA was not detectable in primary cultures of human trophoblast cells incubated for 24 h without or with the cyclic AMP analog. StAR mRNA was also not detected in Northern blots of poly (A)+ RNA isolated from JEG-3 choriocarcinoma cells cultured for 24 h without or with 8-bromo-cAMP (data not shown), a treatment that up-regulates P450scc and adrenodoxin gene expression (Picado-Leonard, J., Voutilainen, R., Kao, L.-C., Chung, B.-C., Strauss, III, J.F., Miller, W.L. (1988) J. Biol. Chem. 263: 3240-3244). These observations suggest that tropic hormones may control levels of StAR in part by increasing the mRNA encoding the protein and hence its synthesis.

Example 6: Mapping of the StAR structural gene and pseudogene

The StAR gene and its pseudogene were mapped by hybridization to Southern blots of DNA from somatic cell hybrids and by polymerase chain reaction analyses using primers specific for the structural gene or pseudogene. High molecular weight genomic DNAs from human x hamster and human x mouse somatic cell hybrid lines obtained from the NIGMS Human Genetic Mutant Cell

Repository (1992/1993 Catalog of Cell Lines, National Institutes of Health) and DNA from human x hamster somatic cell hybrids purchased from BIOS Corporation (New Haven, CT) were used to assign the chromosomal localization of the structural gene and pseudogene.

Regional mapping of the StAR structural gene was accomplished with a chromosome 8 regional mapping panel consisting of hybrids 9HL10, ISHL27 and 20XP0435-2, supplied by Dr. M. Wagner (Chang, Y.J., McCabe, R.T., Rennert, H., Budarf, M.L., Sayegh. R., Emanuel, B.S., Skolnick, P., Strauss, III, J.F. (1992) DNA Cell Biol. 11: 471-480), 8q-, 21q+ and C117 (Wagner, M.J., Ge, Y., Siciliano, M., Wells, D.E. (1991) Genomics 10: 114-125; Dalla-Favera, R., Bregni, M., Erikson, J., Patterson, D., Gallo, R.C., Croce, C.M. (1982) Proc. Natl. Acad. Sci. U.S.A. 82: 464-468; Drabkin, H.A., Diaz, M., Bradley. C.M., Le Beau, MM., Rowley, J.D., Patterson, D. (1985) Proc. Natl. Acad. Sci. U.S.A. 82: 464-468.), and Rec8, which is a hybrid produced by the fusion of the GlyB CHO-K1 mutant with cells from a patient suffering from Recombinant 8 Syndrome (Sacchi, N., Cheng, S.V., Tanzi, R.E., Gusella, J.F., Drabkin, H.A., Patterson, D., Haines, J.H., Papas, T.S. (1988) Genomics 3: 110-116). This cell line contains the Recombinant 8 chromosome, but has no normal human chromosome 8.

When genomic DNA from the hybrid panel was digested with Hind III and subjected to Southern blotting (technical details of Southern blotting are set out below), a strong hybridization band of about 8 kb was detected in the human genomic DNA control and in hybrid GM 10156, which contains only human chromosome 8 (Figure 11). A faint band was also detected in GM 10478, which in addition to containing human chromosome 20 also contains a fragment of human chromosome 8p. These findings indicated that file StAR gene resides on chromosome 8.

To confirm the localization of the StAR gene to chromosome 8, we examined somatic cell hybrid DNA by PCR with primers that specifically amplify the structural gene. Hybrids containing chromosome 8 gave a positive signal

whereas all other hybrids, including those known to contain human chromosome 20 but not 8, did not yield a specific amplification product.

Analysis of a human chromosome 8 regional mapping panel placed the StAR gene on 8p (Figure 12A). Confirmation and refinement of the regional mapping of the functional StAR gene was carried out by isolating a YAC containing the StAR functional gene and using this YAC as a probe in FISH (Figure 12B). Regional mapping was done by sequential banding followed by FISH. By this method the StAR locus was assigned to 8p11.2. Simultaneous FISH with the StAR YAC and an 8 centromere-specific probe as well as fractional length measurements confirmed this assignment.

PCR analysis of reverse transcribed RNA from human testis and PCR analysis of human genomic DNA suggested the existence of an expressed StAR pseudogene. DNA sequences of the amplified pseudogene product did not contain introns and differed in a large number of positions from the functional StAR gene sequence in terms of nucleotide insertions, deletions and substitutions. The amplified sequences differed among several individuals, suggesting significant polymorphism. Using primers specific for the pseudogene sequences, we determined that a StAR pseudogene resides on chromosomes 13 (Figure 13).

Example 7: Southern blotting and PCR Analysis

Ten-12 μg of genomic DNA from each of 24 somatic cell hybrids, total human, hamster (RJK88) and mouse (GM Cl 1-D) were digested with Hind III and electrophoresed through 0.8% agarose and blotted to Hybond N+ (Amersham, Aylesbury, United Kingdom) membranes. Hybridizations with StAR cDNA were performed using previously described conditions (Chang, Y.J., McCabe, R.T., Rennert, H., Budarf, M.L., Sayegh. R., Emanuel, B.S., Skolnick, P., Strauss, III, J.F. (1992) DNA Cell Biol. 11: 471-480).

The StAR structural gene and pseudogene were mapped by PCR analysis of somatic cell hybrid DNA with sequence specific primers. For the structural gene the forward primer used was 5'-GTGAGCAAAGTCCAGGTGCG-3' and the reverse primer was

5'-TGTGGCCATGCCAGCCAGCA-3'. These sequences span a small intron and yield a product of 300 nt. Primers derived from the DNA sequence of the PCR amplified expressed pseudogene, the sequence of which will be reported elsewhere, were used to determine the pseudogene location. The forward primer was 5'-AGCCTCACCGGCGTTGGCGG-3' and the reverse primer was 5'-CTGCAAGACCTTGATCGCCTTG-3'. These primers yield a 800 nt pseudogene-specific product. The PCR conditions were denaturation at 94 C for 5 min followed by a cycle of denaturation at 94 C for 45 sec, annealing at 65 C for 45 sec and extension at 72 C for 2 min for 30 cycles with 10 pM of the primers in a buffer containing 2 mM $MgCl_2$. The PCR products were analyzed by electrophoresis in 1% agarose gels, stained with ethidium bromide.

To confirm the regional mapping of the structural StAR gene, we analyzed the regional mapping panel for several genes known to map to chromosome 8p including the clustrin gene (CL1) (Smith, A.C.M., Spuhler, K., Wiliams, T.M., McConnell, T., Sujansky, E., Robinson, A. (1987) Am. J. Human. Genetics 41: 1083-1103; de Silva, H.V., Harmony, J.A., Stuart, W.D., Gil, C.M., Ribbins, J. (1990) Biochemistry 29: 5380-5389; Jenne, D.E., Tschopp, J. (1989) Proc. Natl. Acad. Sci. U.S.A. 86: 7123-7127; Kirszbaum, L., Sharpe, J.A., Murphy, J., d'Apice, A.J., Classon, B., Hudson, P., Walker, I.D. (1989) EMBO J. 8: 711-718); the lipoprotein lipase gene (LPL) (Pineault, J.M., Tenniswood, M. (1993) J. Biol. Chem. 268: 5021-5031); and the squalene synthase gene (SS) (Wion, K.L., Kirchgessner, T.G., Lusis, A.J., Schotz, M.c., Lawn, R.M. (1987) Science 235: 1638-1641). PCR primers were designed from the published sequences. The CL1-specific primers were 5'-AGAAAGCGCTGCAGGAATACC-3'and5'-GTGACGTGCAGAGCTCTC-3', representing nt 2504-2524 and 2836-2854, respectively. The LPL-specific primers were 5'GAAACTGGGCGAATCTAC-3' and 5'TTGAAACACCCCAAACACTG-3', representing nt 1601-1620 and 1687-1706, respectively. The SS-specific primers were 5'-AAAAGAACGCTGTGTGGCIGGGAC-3' and

5'-ACCTAAACCGTGGCAAAT-3', representing nt 1405-1428 and 1547-1568, respectively.

Example 8: Fluorescence *in situ* hybridization (FISH) mapping

An individual yeast artificial chromosome (YAC) colony containing the StAR structural gene was isolated from the St. Louis library by PCR screening using StAR-specific primers corresponding to the 3'-untranslated sequences. The sense primer was 5'-CCTACTGGAAGCCTGCAAGTCTAAG-3' (nt 1048-1072). The antisense primer was 5'-TGGTTTTAGGTGGGTACATAAGGG-3' (nt 1287-1264). StAR sequences in YAC DNA were amplified in a standard PCR reaction vol of 10 μl containing 1 mM $MgCl_2$. YAC DNA was initially denatured at 94 C for 5 min. Amplification was carried out with 35 cycles of denaturation at 94 C for 30 sec, annealing at 55 C for 30 sec and extension at 72 C for 30 sec. The reaction products were analyzed for the presence of the expected 240 nt amplification product in 2% agarose gels followed by ethidium bromide staining.

YAC FISH was performed as previously described (Jiang, G., McKenzie, T.L., Conrad, D.G., Schechter, I. (1993) J. Biol. Chem. 268: 12818-12824; Lichter, P., Tang, C.-J. C., Call, K., Hermanson, G., Glen, A.E., Housman, D., Ward, D.C. (1990) Science 247: 64-69) with the following modifications: The biotin-labeled probe was denatured at 75 C for 5 min, pre-annealed with human Cot-1 DNA for 1 h at 37 C and applied to human chromosome slide preparations that had been previously denatured and dehydrated. Slides were cover-slipped and hybridized overnight in a humid chamber at 37 C. In some experiments, a chromosome 8 centromere-specific probe (D8Z2; Oncor, Inc., Gaithersburg, MD) was added to the hybridization mixture. Post-hybridization washes were done in 50% formamide/2 X SSC (1 X SSC=O.15 M NaCl and 0.015 M sodium citrate) for 15 min and 2 X SSC for 8 min. at 45 C. Detection was by avidin-FITC, with one amplification by the manufacturer's directions (Oncor, Inc.). Chromosomes were counter-stained with propidium iodide.

Twelve metaphase spreads were G-banded by trypsin and photographed prior to FISH. Slides were washed in Heme-De (Fisher Scientific, Fairlawn, NJ) to remove the oil, destained in absolute methanol two times for 10 min, dehydrated in 70% and then 80% ethanol for 2 min each, placed in absolute methanol for 10 min and air dried. FISH was then performed as described above. Metaphase spreads were relocated and banding patterns compared with probe signal to assign location of the probe. Fractional length measurements confirmed the assignment (Jiang, G., McKenzie, T.L., Conrad, D.G., Schechter, I. (1993) J. Biol. Chem. 268: 12818-12824).

Metaphase spreads were either photographed with a Zeiss Axiophot microscope with Ektachrome 400 slide film, or processed digitally by computer and printed with a color printer.

Example 9: RNase protection of Native RNA from Patient 5 and RNAs produced *in vitro*

To determine the effect of the T → A transversion on RNA splicing we constructed a pCMV4 expression vector for a human StAR minigene. This vector expresses a primary RNA transcript consisting Exons 1, 2, 3, 4, Intron 4, Exon 5, Intron 5, Exon 6, Intron 6, and Exon 7; i.e. the first four exons are from the cDNA and the remainder of the construct is from the native gene.

Minigene constructs - The normal and mutant minigene constructs contain 5'-exons 1-2-3-4, intron 4, exon 5, intron 5, exon 6, intron 6, exon 7, 3', cloned into pCMV4. The cDNA portion (exons 1-2-3-4) was produced by amplification with 5'-ATACTAAGCTTGCAACCACCCTTGAGAGAAG (bases 41-60), and 5'-CCCCATTGTCCTGCTGACTCTC (bases 421-442); the genomic portion (exon 4, intron 4, exon 5, intron 5, exon 6, intron 6, exon 7) was amplified using 5'-GGGGACAAAGTGATGAGTAAAGTG (bases 439-462), and 5'-ATATCTAGACTGATGAGCGTGTGTACCAG (bases 1021-1040). Following PCR, excess nucleotides were removed by centrifugation through Centricon-100. DNA fragments were treated with the Klenow fragment of DNA polymerase I in the present of 50 μM dTTP, and cloned downstream from the CMV promoter.

Five micrograms of each construct were transfected by calcium phosphate precipitation into COS-1 cells grown in DMEM-H21 supplemented with 10% fetal calf serum. An RSV-LUC construct (0.5 μg) was used to monitor transfection efficiency in each plate. The media were changed 12 hours after transfection, and cells were grown for another 48 hours before harvesting. Luciferase activity was measured using the Luciferase Assay System (Promega) for 15 secs. Total RNA was prepared by guanidine thiocyanate extraction, while cytoplasmic and nuclear RNA were prepared by cell lysis with NP40.

The splicing of this RNA was then assessed by RNase protection assays using riboprobes consisting of exons 4 and 5 (probe 1), or exon 4, intron 4, exon 5 and exon 6 (probe 2). DNA fragments were amplified using 5'-ATAGAATTCGACAAAGTGATGAGTAAAGTG (bases 442-462) as the sense primer (S5) and the anti-sense primer (AS6) 5'-CTGATGACACCCTTCTGCTC (bases 757-776) for the construct for the first probe and the antisense primer (AS7) 5'-CTTGAGGTCGATGCTGAGTAGCCTAGGATA (bases 850-870) for the construct for the second probe. These fragments were cloned into *Eco* RI and *Bam* HI sites of pKS. The construct for the third probe was produced from ligation of a *Pst* I/*Eco* RI genomic fragment to a *Pst* I/*Eco* RI fragment of the second construct. RNase protection experiments were done as described (*Mol. Cell Biol.* 12, 2124-2134 (1992)). Probe 1 detects a single 335 bp band from cells transfected with the normal minigene vector, corresponding to exons 4 and 5 without any intervening intronic sequences (Figure 14). However in cells transfected with the vector containing the intron 4 T → A transversion, no 335 bp band is seen but instead exon 5 (185 bp) and exon 4 (150 bp) are protected as separate fragments, indicating that there were additional RNA sequences separating exons 4 and 5, presumably unspliced intron 4. Exon 4 consistently yielded a stuttered protection pattern; the basis of this is unknown.

If the T → A transversion in intron 4 causes the splicing machinery to skip exon 5, as suggested by the RT-PCR data from the patient's gonad, then we would expect to find exon 4 fused to exon 6 in the RNA of cells transfected with the mutant minigene construct, but not in the RNA of cells transfected with the

wild-type minigene construct. To assess this, we performed and RNase protection assays with probe 2 (Figure 15). This probe detects fragments of 150 bases, corresponding to exon 4, and of 279 bases corresponding to exons 5 and 6, in control cDNA and in the RNA from cells transfected with either the normal or mutant constructs. The RNA from the transfected cells also protects unspliced fragments of 570 bp (exon 4, intron 4, exons 5 and 6) and of 476 bp (exon 4, intron 4, exon 5). The 150 and 279 bp spliced forms are much more abundant than the unspliced forms expressed by the normal construct, but are of equivalent intensity from the mutant construct. Thus the T → A transversion interferes with normal splicing. When nuclear and cytoplasmic RNAs are separated, it is apparent that most of the intronic species are nuclear, but there are some intronic species in the cytoplasm and some spliced forms in the nucleus, reflecting either cellular leakage or contamination of each fraction with the other during cell fractionation. It is also possible that very high levels of expression from the pCMV4 vector might saturate the splicing machinery. Thus analysis of minigene expression constructs indicates that a single T → A transversion 11 bp from the intron 4/exon 5 splice site disrupts proper splicing of StAR pre-RNA, causing lipoid CAH.

RNase protection of native RNA - A small amount of the patient's testicular RNA was available for direct examination with probe 2 (Figure 16). Normal human fetal adrenal RNA only protected the expected fragments of 279 bases, corresponding to exons 5 and 6, and of 150 bases, corresponding to exon 4, as seen previously in (Figure 14). Thus the PCR products that lack exon 5, which occasionally were seen with normal RNA, represented very rare events. The patient's RNA protected several additional species. There were small amounts of 570 and 476 base species, corresponding to the retention of unspliced intron 4 associated with exons 4, 5 and 6 (570 bp) or associated only with exons 4 and 5 (476 bp). The abundance of the 150 base species was only slightly less than in the normal RNA, but the abundance of the 279 base species was much less, consistent with the T → A transversion causing skipping of exon 5 in much of the patient's RNA. The patient's RNA also had a large amount of a species running close to the 100 bp DNA marker, presumably corresponding to exon 6 (94 bases), which

would be expected to be protected but not joined to other sequences if the T → A transversion causes skipping of exon 5. Finally, the patient's RNA contains a small amount of the 185 bp species corresponding to isolated exon 5, unassociated with exon 6. Thus the T → A transversion appears to cause several forms of disordered RNA splicing: skipping of exon 5 as seen in the PCR-amplified cDNA, is the predominant error, but retention of intron 4 and additional downstream splicing errors, reflected by the 185 base species, also occur.

Example 10: Transfection Studies and Analysis of 3β-Hydroxy-5-cholestenoic Acid.

COS-1 cells were transfected with the rat P450c27 cDNA in pCMV4 (Su et al., (1990)) and a bovine adrenodoxin expression plasmid kindly provided by Dr. Michael Waterman (Vanderbilt University) with either the pSV-SPORT-1 empty vector (BRL, Bathesda, MD) or the vector containing the human StAR cDNA as previously described (Sugawara et al., PNAS 1995).

Formation of 3β-hydroxy-5-cholestenoic acid, the end-product of cholesterol metabolism by P450c27 (Andersson et al., 1989), was analyzed by isotope mass spectrometry as previously described (Reiss et al., *J. Lipid Res.* **35**, 1026-1030 (1994)). In brief, deuterated standard (500 ng) was added to 1 mL aliquots of medium, and after acidification and extraction into ethyl acetate the product was isolated by thin-layer chromatography. After methylation of the C_{27} acid, the eluates were acetylated and injected into a Hewlett-Packard GLC—MS onto a fused silica column (CP-sil 19CB, 0.25 mm i.d., 25 m length; Chrompack, Raritan, NJ). Using an isotope ratio program, the areas for *m/z* 4/4 [3β-hydroxy-5-cholestenoic acid methyl ester acetate, molecular ion = 472 - 60 (acetate)] were monitored and the respective areas determined by integration for calculation of the amount of endogenous sterol. The results are shown in Table 12.

Table 12

Effect of StAR on Mitochondrial Cholesterol 27-Hydroxylase Activity[a]

group	plasmid			(pmol/mL) 3β-hydroxy-5-cholestenoic acid[b]
	P450c27	StAR	adrenodoxin	
1	-	-	+	0.90 ± 0.42 (6)
2	-	+	+	0.44 ± 0.32 (5)
3	+	-	+	3.25 ± 0.70 (6)
4	+	+	+	20.64 ± 3.79 (6)

[a] COS-1 cells were transfected with the indicated expression plasmids for bovine adrenodoxin, StAR, or the control vector pSV-SPORT-1, introduced when the StAR plasmid was not included, described (Sugawara et al., 1995). Media were collected 48 h after transfection for quantification of 3β-hydroxy-5-cholestenoic acid.

[b] Values presented are means ± SE of determinations from the indicated number of dishes from three separate experiments.

Example 11: Expression of a Heterologous Gene Using StAR Promoter

To analyze the promoter activity of the StAR promoter, a 1.3 kb *Hind*III fragment of the StAR gene (bp — 1293 to +25) was cloned in the correct and reverse orientation into the plasmid vector pGL_2 (Promega) which contains firefly luciferase as a reporter gene. Other plasmids used in these experiments included the pGL_2 basic vector, which contains no promoter sequences; pGL_2 control, which places the luciferase gene under the control of the SV40 promoter and enhancer; and pCH110, a plasmid in which the *lac*Z gene is under control of the early SV40 promoter (Pharmacia).

Murine Y-1 adrenal tumor cells and BeWo choriocarcinoma cells were grown in 35 mm plastic dishes in a culture medium consisting of Dulbecco's minimal essential medium supplemented with 10% fetal calf serum and 50 μg/mL of gentamycin. Plasmids used for transfection were purified using the Maxiprep reagent system (Qiagen). Cell cultures at 40%-60% confluence were washed twice with serum-free medium before adding 1 mL of serum-free medium containing 1 μg of pGL_2 plasmid constructs and 1 μg of pCH110 plasmid with 10 μL of

Lipofectamine (GIBCO/BRL). After 5 h of incubation, the medium was replaced with 1 mL of medium with 20% serum. Cells were harvested after 48 h of culture. In some experiments, 8-Br-cAMP (1 mM) was added to the medium for the final 24 h of culture.

Cells were harvested 48 h after transfection, and extracts were made in Promega lysis buffer. One aliquot (40 μL out of 400 μL total extract volume) was used for luciferase assays with Promega reagents, and another 150 μL was taken for β-galactosidase assays with Promega reagents. The "blank" luciferase value measured in untransfected cell extracts was subtracted from luciferase readings of transfected cell extracts. The luciferase assay results were normalized to β-galactosidase activity to compensate for variations in transfection efficiency. In each experiment the activity of the pGL_2 control vector was defined as 100%. Each treatment group contained at least triplicate cultures, and each experiment was repeated two or three times. The results are shown in Table 13.

Table 13

StAR Gene Promoter Activity in Y-1 Adrenal Cortical Tumor Cells and BeWo Choriocarcinoma Cells[a]

plasmid/treatment	Y-1 cells (%)	BeWo cells (%)
pGL_2 control	100	100
pGL_2 control + cAMP	109 ± 6	118 ± 14
pGL_2 basic	0.8 ± 0.1	0.17 ± 0.02
pGL_2 basic + cAMP	1.1 ± 0.9	0.4 ± 0.04
pGLStAR 1.3 kb	17.8 ± 4	0.8 ± 0.1
pGLStAR 1.3 kb + cAMP	42.8 ± 8	1.5 ± 0.1

The results presented are the mean ± SE of four separate transfections carried out with triplicate cultures for each treatment for Y-1 cells and three separate transfections carried out with triplicate cultures for BeWo cells.

Treatment of the Y-1 cells with 8-Br-cAMP increased StAR promoter activity 2.3-fold ($p < 0.002$ analysis of log-transformed data by the paired *t*-test), suggesting that this segment of DNA contains cAMP-responsive elements. This increase in StAR promoter activity corresponds to the 3-fold increase in steady state levels of StAR mRNA in human granulosa cells treated for 4 h with cAMP (Sugawara et al., (1995)). This suggests that the increase in StAR mRNA in response to cAMP is, at least in part, the result of increased transcription.

In contrast to our findings with Y-1 cells, the StAR promoter did not cause significant reporter gene expression in BeWo choriocarcinoma cells, which do not express the StAR gene. This was true whether BeWo cells were examined in the basal state or after stimulation with 8-Br-cAMP. This is consistent with the lack of detectable StAR mRNA in placenta and choriocarcinoma cells (Sugawara et al., (1995)) and with the persistence of placental steroidogenesis in pregnancies in which the fetus is affected with lipoid CAH (Saenger et al., 1995). Thus cis

elements directing the tissue-specific expression of StAR and regulation by cAMP appear to be located within 1.3 kb of DNA upstream from the cap site.

All publications and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference.

The invention now being fully described, it will be apparent to one of ordinary skill in the art that many changes and modifications can be made thereto without departing from the spirit or scope of the appended claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANTS: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA; THE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA

(ii) TITLE OF INVENTION: IDENTIFICATION OF GENE MUTATIONS ASSOCIATED WITH CONGENITAL LIPOID ADRENAL HYPERPLASIA

(iii) NUMBER OF SEQUENCES: 30

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Robbins, Berliner & Carson
(B) STREET: 201 W. Figueroa Street, 5th Floor
(C) CITY: Los Angeles
(D) STATE: California
(E) COUNTRY: US
(F) ZIP: 90012-2628

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Berliner, Robert
(B) REGISTRATION NUMBER: 20,121
(C) REFERENCE/DOCKET NUMBER: 5555-366C1

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 213-977-1001
(B) TELEFAX: 213-977-1003
(C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1618 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 127..984

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
CCACGCGTCC GCGAAGCTTG AGGGGCTCAG GAAGGACGAA GCAACCACCC TTGAGAGAAG      60

AGGCAGCAGC AGCGGCGGCA GCAGCAGCGG CAGCGACCCC ACCACTGCCA CATTTGCCAG     120

GAAACA ATG CTG CTA GCG ACA TTC AAG CTG TGC GCT GGG AGC TCC TAC         168
       Met Leu Leu Ala Thr Phe Lys Leu Cys Ala Gly Ser Ser Tyr
         1               5                  10

AGA CAC ATG CGC AAC ATG AAG GGG CTG AGG CAA CAG GCT GTG ATG GCC       216
Arg His Met Arg Asn Met Lys Gly Leu Arg Gln Gln Ala Val Met Ala
 15                  20                  25                  30

ATC AGC CAG GAG CTG AAC CGG AGG GCC CTG GGG GGC CCC ACC CCT AGC       264
Ile Ser Gln Glu Leu Asn Arg Arg Ala Leu Gly Gly Pro Thr Pro Ser
                 35                  40                  45

ACG TGG ATT AAC CAG GTT CGG CGG CGG AGC TCT CTA CTC GGT TCT CGG       312
Thr Trp Ile Asn Gln Val Arg Arg Arg Ser Ser Leu Leu Gly Ser Arg
            50                  55                  60

CTG GAA GAG ACT CTC TAC AGT GAC CAG GAG CTG GCC TAT CTC CAG CAG       360
Leu Glu Glu Thr Leu Tyr Ser Asp Gln Glu Leu Ala Tyr Leu Gln Gln
        65                  70                  75

GGG GAG GAG GCC ATG CAG AAG GCC TTG GGC ATC CTT AGC AAC CAA GAG       408
Gly Glu Glu Ala Met Gln Lys Ala Leu Gly Ile Leu Ser Asn Gln Glu
    80                  85                  90

GGC TGG AAG AAG GAG AGT CAG CAG GAC AAT GGG GAC AAA GTG ATG AGT       456
Gly Trp Lys Lys Glu Ser Gln Gln Asp Asn Gly Asp Lys Val Met Ser
 95                  100                 105                 110

AAA GTG GTC CCA GAT GTG GGC AAG GTG TTC CGG CTG GAG GTC GTG GTG       504
Lys Val Val Pro Asp Val Gly Lys Val Phe Arg Leu Glu Val Val Val
                115                 120                 125

GAC CAG CCC ATG GAG AGG CTC TAT GAA GAG CTC GTG GAG CGC ATG GAA       552
Asp Gln Pro Met Glu Arg Leu Tyr Glu Glu Leu Val Glu Arg Met Glu
            130                 135                 140

GCA ATG GGG GAG TGG AAC CCC AAT GTC AAG GAG ATC AAG GTC CTG CAG       600
Ala Met Gly Glu Trp Asn Pro Asn Val Lys Glu Ile Lys Val Leu Gln
        145                 150                 155

AAG ATC GGA AAA GAT ACA TTC ATT ACT CAC GAG CTG GCT GCC GAG GCA       648
Lys Ile Gly Lys Asp Thr Phe Ile Thr His Glu Leu Ala Ala Glu Ala
    160                 165                 170

GCA GGA AAC CTG GTG GGG CCC CGT GAC TTT GTG AGC GTG CGC TGT GCC       696
Ala Gly Asn Leu Val Gly Pro Arg Asp Phe Val Ser Val Arg Cys Ala
175                 180                 185                 190

AAG CGC CGA GGC TCC ACC TGT GTG CTG GCT GGC ATG GAC ACA GAC TTC       744
Lys Arg Arg Gly Ser Thr Cys Val Leu Ala Gly Met Asp Thr Asp Phe
                195                 200                 205

GGG AAC ATG CCT GAG CAG AAG GGT GTC ATC AGG GCG GAG CAC GGT CCC       792
Gly Asn Met Pro Glu Gln Lys Gly Val Ile Arg Ala Glu His Gly Pro
            210                 215                 220

ACT TGC ATG GTG CTT CAC CCG TTG GCT GGA AGT CCC TCT AAG ACC AAA       840
Thr Cys Met Val Leu His Pro Leu Ala Gly Ser Pro Ser Lys Thr Lys
        225                 230                 235

CTT ACG TGG CTA CTC AGC ATC GAC CTC AAG GGG TGG CTG CCC AAG AGC       888
Leu Thr Trp Leu Leu Ser Ile Asp Leu Lys Gly Trp Leu Pro Lys Ser
    240                 245                 250
```

ATC ATC AAC CAG GTC CTG TCC CAG ACC CAG GTG GAT TTT GCC AAC CAC 936
Ile Ile Asn Gln Val Leu Ser Gln Thr Gln Val Asp Phe Ala Asn His
255 260 265 270

CTG CGC AAG CGC CTG GAG TCC CAC CCT GCC TCT GAA GCC AGG TGT TGAAGACCAG 991
Leu Arg Lys Arg Leu Glu Ser His Pro Ala Ser Glu Ala Arg Cys
275 280 285

CCTGCTGTTC CCAACTGTGC CCAGCTGCAC TGGTACACAC GCTCATCAGG AGAATCCCTA 1051

CTGGAAGCCT GCAAGTCTAA GATCTCCATC TGGTGACAGT GGGATGGGTG GGGTTCGTGT 1111

TTAGAGTATG ACACTAGGAT TCAGATTGGT GAAGTTTTTA GTACCAAGAA AACAGGGATG 1171

AGGCTCTTGG ATTAAAAGGT AACTTCATTC ACTGATTAGC TATGACATGA GGGTTCAGGC 1231

CCCTAAAATA ATTGTAAAAC TTTTTTTCTG GGCCCTTATG TACCCACCTA AAACCATCTT 1291

TAAAATGCTA GTGGCTGATA TGGGTGTGGG GGATGCTAAC CACAGGGCCT GAGAAGTCTT 1351

GCTTTATGGG CTCAAGAATG CCATGCGCTG GCAGTACATG TGCACAAAGC AGAATCTCAG 1411

AGGGTCTCCT GCAGCCCTCT GCTCCTCCCG GCCGCTGCAC AGCAACACCA CAGAACAAGC 1471

AGCACCCCAC AGTGGGTGCC TTCCAGAAAT ATAGTCCAAG CTTTCTCTGT GGAAAAAGAC 1531

AAAACTCATT AGTAGACATG TTTCCCTATT GCTTTCATAG GCACCAGTCA GAATAAAGAA 1591

TCATAATTCA CACCAAAAAA AAAAAAA 1618

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 285 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Leu Leu Ala Thr Phe Lys Leu Cys Ala Gly Ser Ser Tyr Arg His
1 5 10 15

Met Arg Asn Met Lys Gly Leu Arg Gln Gln Ala Val Met Ala Ile Ser
20 25 30

Gln Glu Leu Asn Arg Arg Ala Leu Gly Gly Pro Thr Pro Ser Thr Trp
35 40 45

Ile Asn Gln Val Arg Arg Arg Ser Ser Leu Leu Gly Ser Arg Leu Glu
50 55 60

Glu Thr Leu Tyr Ser Asp Gln Glu Leu Ala Tyr Leu Gln Gln Gly Glu
65 70 75 80

Glu Ala Met Gln Lys Ala Leu Gly Ile Leu Ser Asn Gln Glu Gly Trp
85 90 95

Lys Lys Glu Ser Gln Gln Asp Asn Gly Asp Lys Val Met Ser Lys Val
100 105 110

Val Pro Asp Val Gly Lys Val Phe Arg Leu Glu Val Val Val Asp Gln
115 120 125

Pro Met Glu Arg Leu Tyr Glu Glu Leu Val Glu Arg Met Glu Ala Met
130 135 140

Gly Glu Trp Asn Pro Asn Val Lys Glu Ile Lys Val Leu Gln Lys Ile
145 150 155 160

Gly Lys Asp Thr Phe Ile Thr His Glu Leu Ala Ala Glu Ala Ala Gly
165 170 175

Asn Leu Val Gly Pro Arg Asp Phe Val Ser Val Arg Cys Ala Lys Arg
180 185 190

Arg Gly Ser Thr Cys Val Leu Ala Gly Met Asp Thr Asp Phe Gly Asn
195 200 205

Met Pro Glu Gln Lys Gly Val Ile Arg Ala Glu His Gly Pro Thr Cys
210 215 220

Met Val Leu His Pro Leu Ala Gly Ser Pro Ser Lys Thr Lys Leu Thr
225 230 235 240

Trp Leu Leu Ser Ile Asp Leu Lys Gly Trp Leu Pro Lys Ser Ile Ile
245 250 255

Asn Gln Val Leu Ser Gln Thr Gln Val Asp Phe Ala Asn His Leu Arg
260 265 270

Lys Arg Leu Glu Ser His Pro Ala Ser Glu Ala Arg Cys
275 280 285

WHAT IS CLAIMED IS:

1. An isolated DNA or RNA molecule, wherein said molecule contains:

(1) a first sequence consisting of hStAR cDNA, hStAR genomic DNA, or hStAR pseudogene DNA as set forth in Figure 1, Table 6, or Table 7;

(2) a second sequence, wherein said second sequence is a subsequence of said first sequence at least 10 nucleotides in length;

(3) a third sequence in which at least one nucleotide of said first or second sequence is replaced by a different nucleotide;

(4) a fourth sequence in which at least one nucleotide is deleted from or inserted into said first or second sequence; or

(5) a fifth sequence complementary to any of said first second, or third sequences;

with the provisos that (1) said molecule can be an RNA molecule in which U replaces T in any of said sequences (1) - (5), (2) said third sequence is at least 95% identical to said first or second sequence, (3) said second sequence is not present in mouse StAR cDNA, and (4) said fourth sequence contains no more than 20 inserted nucleotides and no more than 200 deleted nucleotides.

2. An isolated DNA molecule, wherein said molecule contains:

(1) a first sequence, consisting of hStAR cDNA, hStAR genomic DNA, or hStAR pseudogene DNA as set forth in Figure 1, Table 6, or Table 7;

(2) a second sequence, wherein said second sequence is a subsequence of said first sequence at least 10 nucleotides in length;

(3) a third sequence in which at least one nucleotide within exon 1, exon 2, exon 3, exon 4, exon 5, exon 6, exon 7, or intron 4 of said first or second sequence is replaced by a different nucleotide; or is deleted from or inserted into said first or second sequence;

(4) a fourth sequence in which at least one nucleotide within exon 1, exon 2, exon 3, exon 4, exon 5, exon 6, exon 7, or intron 4 of said first or second sequence is deleted from or inserted into said first or second sequence; or

(4) a fifth sequence complementary to any of said first second, or third sequences;

with the provisos that (1) said molecule can be an RNA molecule in which U replaces T in any of said sequences (1) - (5), (2) said third sequence is at least 95% identical to said first or second sequence, (3) said second sequence is not present in mouse StAR cDNA, and (4) said fourth sequence contains no more than 20 inserted nucleotides and no more than 200 deleted nucleotides.

3. The isolated molecule of Claim 1, wherein said molecule comprises said first sequence.

4. The isolated molecule of Claim 1, wherein said molecule comprises said second sequence.

5. The isolated molecule of Claim 1, wherein said molecule comprises said third or fourth sequence.

6. The isolated molecule of Claim 5, wherein said different nucleotide, said deleted nucleotide, or said inserted nucleotide is present in an exon.

7. The isolated molecule of Claim 6, wherein said different nucleotide, said deleted nucleotide, or said inserted nucleotide is present in exon 5, exon 6, or exon 7.

8. The isolated molecule of Claim 5, wherein said different nucleotide, said deleted nucleotide, or said inserted nucleotide is present in intron 4.

9. The isolated molecule of Claim 8, wherein said different nucleotide is a T→A transversion 11 bp from the junction of intron 4 and exon 5.

10. The isolated molecule of Claim 5, wherein said different nucleotide causes the codon in which said nucleotide is located to be a TAA, TGA, or TAG stop codon.

11. The isolated molecule of Claim 6, wherein said different nucleotide is an *Arg*^{193}→*Stop* mutation or a *Gln*^{258}→*Stop* mutation.

12. The isolated molecule of Claim 6, wherein said different nucleotide causes an amino acid replacement.

13. The isolated molecule of Claim 12, wherein said different nucleotide causes a *Glu*^{169}→*Gly* replacement, an *Arg*^{182}→*Leu* replacement, a *Glu*^{169}→*Lys* replacement, an *Ala*^{218}→*Val* replacement, or a *Leu*^{275}→*Pro* replacement.

14. The isolated molecule of Claim 6, wherein said deleted nucleotide is T^{593}.

15. The isolated molecule of Claim 6, wherein said inserted nucleotide is a G between G^{247} and G^{248} or an A between G^{947} and C^{948}.

16. The isolated molecule of Claim 6, wherein said deleted nucleotides are C^{940}, G^{941}, and C^{942}.

17. The isolated molecule of Claim 5, wherein said different nucleotide, said deleted nucleotide, or said inserted nucleotide inhibits, directly or indirectly, transcription of said steroidogenesis acute regulatory protein gene.

18. The isolated molecule of Claim 5, wherein said different nucleotide, said deleted nucleotide, or said inserted nucleotide inhibits, directly or indirectly, translation of mRNA of said steroidogenesis acute regulatory protein gene.

19. The isolated molecule of Claim 5, wherein said different nucleotide, said deleted nucleotide, or said inserted nucleotide decreases, directly or indirectly, stability of mRNA of said steroidogenesis acute regulatory protein.

20. A method of detecting the presence of a genetic defect that has the potential of causing congenital lipoid adrenal hyperplasia in a human or of transmitting congenital lipoid adrenal hyperplasia to an offspring of said human, which comprises:

obtaining nucleic acid containing a gene encoding a steroidogenesis acute regulatory protein from said human;

analyzing said nucleic acid for the presence or absence of a mutation of said gene, wherein said mutation provides a sequence different from human steroidogenesis acute regulatory protein genomic DNA sequence as set forth in Table 6, whereby presence of said mutation is indicative of a genetic defect having a potential of causing congenital lipoid adrenal hyperplasia.

21. A method of detecting the presence of a genetic defect that causes congenital lipoid adrenal hyperplasia in a human or that can transmit congenital lipoid adrenal hyperplasia to an offspring of said human, which comprises:

obtaining nucleic acid containing a gene encoding a steroidogenesis acute regulatory protein from said human;

analyzing said nucleic acid for the presence or absence of a mutation of said gene, wherein said mutation provides a sequence different from human steroidogenesis acute regulatory protein genomic DNA sequence as set forth in Table 6 and said mutation is known to be indicative of a genetic defect that causes congenital lipoid adrenal hyperplasia when present in said human in the absence of a heterozygous steroidogenesis acute regulatory protein gene have a genomic DNA sequence as set forth in Figure 10.

22. A method of detecting the presence of a genetic defect that has the potential of causing congenital lipoid adrenal hyperplasia in a human or of transmitting

congenital lipoid adrenal hyperplasia to an offspring of said human, which comprises:

obtaining nucleic acid containing a gene encoding a steroidogenesis acute regulatory protein from said human;

analyzing said nucleic acid for the presence or absence of a mutation of said gene, wherein said mutation provides a sequence different from human steroidogenesis acute regulatory protein genomic DNA sequence of exon 1, exon 2, exon 3, exon 4, exon 5, exon 6, exon 7, or intron 4, whereby presence of said mutation is indicative of a genetic defect having a potential of causing congenital lipoid adrenal hyperplasia.

23. The method of Claim 20, wherein said mutation results in a change in the sequence of a protein product of said steroidogenesis acute regulatory protein gene.

24. The method of Claim 20, wherein said mutation results in said steroidogenesis acute regulatory protein gene not being transcribed or translated.

25. The method of Claim 20, wherein said mutation creates a stop codon in said steroidogenesis acute regulatory protein gene.

26. The method of Claim 25, wherein said mutation is an $Arg^{193} \rightarrow Stop$ mutation or a $Gln^{258} \rightarrow Stop$ mutation.

27. The method of Claim 20, wherein said method comprises PCR amplification of at least a segment of said steroidogenesis acute regulatory protein gene.

28. The method of Claim 20, wherein said method comprises identifying a change in a restriction site as a result of said mutation.

29. The method of Claim 20, wherein said method comprises restriction fragment length polymorphism analysis, allele-specific oligonucleotide hybridization, or nucleotide sequencing.

30. The method of Claim 20, wherein said method classifies said human as homozygous for said steroidogenesis acute regulatory protein gene or for said mutated steroidogenesis acute regulatory protein gene or heterozygous for said steroidogenesis acute regulatory protein gene and said mutated steroidogenesis acute regulatory protein gene.

31. The method of Claim 23, wherein said mutation is a $Glu^{169} \rightarrow Gly$ mutation, $Arg^{182} \rightarrow Leu$ mutation, a $Glu^{169} \rightarrow Lys$ mutation, an $Ala^{218} \rightarrow Val$ mutation, $Leu^{275} \rightarrow Pro$ mutation.

32. The method of Claim 20, wherein said mutation is a deletion or insertion mutation, said deletion or insertion mutation comprising at least one deleted nucleotide, at least one inserted nucleotide, or at least one inserted and at least one deleted nucleotide.

33. The method of Claim 32, wherein said mutation is a frame-shift mutation.

34. The method of Claim 33, wherein said deleted nucleotide is T^{593} or C^{650}.

35. The method of Claim 33, wherein said inserted nucleotide is a G between G^{247} and G^{248} or an A between G^{947} and C^{948}.

36. The method of Claim 32, wherein said deleted nucleotides are C^{940}, G^{941}, and C^{942}.

37. The method of Claim 20, wherein said mutation causes aberrant splicing of the mRNA encoded by the steroidogenesis acute regulatory protein gene.

38. The method of Claim 20, wherein said mutation inhibits transcription of said steroidogenesis acute regulatory protein gene.

39. The method of Claim 20, wherein said mutation inhibits translation of mRNA of said steroidogenesis acute regulatory protein.

40. The method of Claim 20, wherein said mutation decreases stability of mRNA of said steroidogenesis acute regulatory protein.

41. The method of Claim 28, wherein said restriction site is selected from the group consisting of Sau 96I, Fsp I, Hha I, Hae II, Nco I, Alu I, Tsp 45I, Ava II, Hae III, Eco RII.

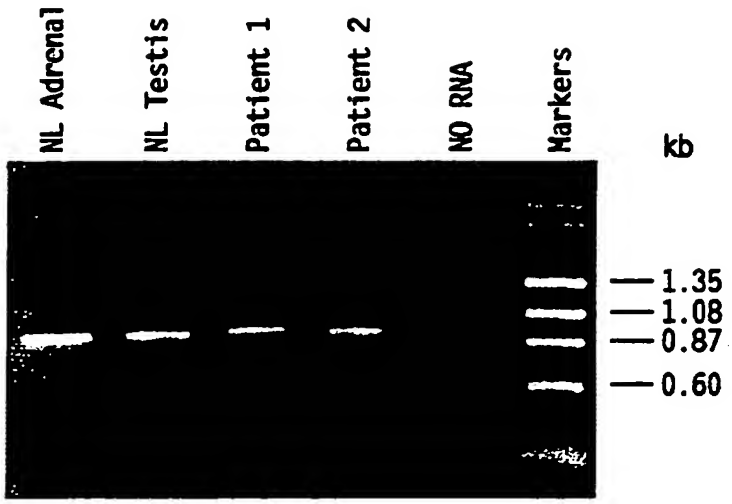
FIG. 1(a)

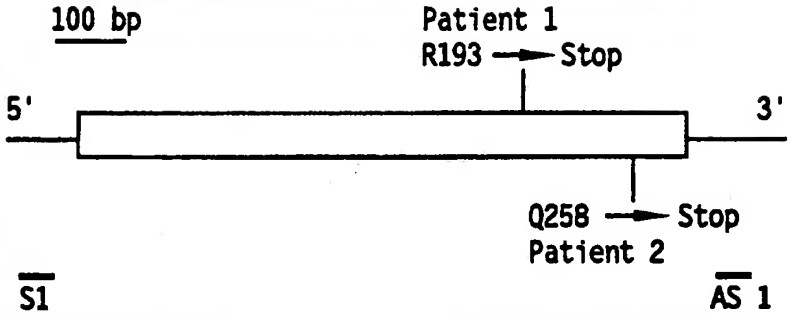
```
                    20                  40                  60                  80                  100                 120
CCA CGC GTC CGC GAA GCT TGA GGG GCT CAG GAA GGA CGA AGC AAC CAC CCT TGA GAG AAG AGG CAG CAG CAG CGG CGG CAC CAC CAG CGG CAG CCC CAC CAC TGC CAC ATT TGC CAG
                   140                 160                 180                 200                 220                 240
GAA ACA ATG CTG CTA GCG ACA TTC AAG CTG TGC GCT GGG AGC TCC TAC AGA CAC ATG CGC AAC ATG AAG GGG CTG AGG CAA CAG GCT GTG ATG GCC ATC AGC CAG CAG CTG AAC CGG AGG
      Met Leu Leu Ala Thr Phe Lys Leu Cys Ala Gly Ser Ser Tyr Arg His Met Arg Asn Met Lys Gly Leu Arg Gln Gln Ala Val Met Ala Ile Ser Gln Glu Leu Asn Arg Arg>
                   260                 280                 300                 320                 340                 360
GCC CTG GGG GGC CCC ACC CCT AGC ACG TGG ATT AAC CAG GTT CGG CGG CGG AGC TCT CTA CTC GGT TCT CGG CTG GAA GAG ACT CTC TAC AGT GAC CAG GAG CTG GCC TAT CTC CAG CAG
Ala Leu Gly Gly Pro Thr Pro Ser Thr Trp Ile Asn Gln Val Arg Arg Arg Ser Ser Leu Leu Gly Ser Arg Leu Glu Glu Thr Leu Tyr Ser Asp Gln Glu Leu Ala Tyr Leu Gln Gln>
                   380                 400                 420                 440                 460                 480
GGG GAG GAG GCC ATG CAG AAG GCC TTG GGC ATC CTT AGC AAC CAA GAG GGC TGG AAG AAG GAG AGT CAG CAG GAC AAT GGG GAC AAA GTG ATG AGT AAA GTG GTC CCA GAT GTG GGC AAG
Gly Glu Glu Ala Met Gln Lys Ala Leu Gly Ile Leu Ser Asn Gln Glu Gly Trp Lys Lys Glu Ser Gln Gln Asp Asn Gly Asp Lys Val Met Ser Lys Val Val Pro Asp Val Gly Lys>
                   500                 520                 540                 560                 580                 600
GTG TTC CGG CTG GAG GTC GTG GTG GAC CAG CCC ATG GAG AGG CTC TAT GAA GAG CTC GTG GAG CGC ATG GAA GCA ATG GGG GAG TGG AAC CCC AAT GTC AAG GAG ATC AAG GTC CTG CAG
Val Phe Arg Leu Glu Val Val Val Asp Gln Pro Met Glu Arg Leu Tyr Glu Glu Leu Val Glu Arg Met Glu Ala Met Gly Glu Trp Asn Pro Asn Val Lys Glu Ile Lys Val Leu Gln>
                   620                 640                 660                 680                 700                 720
AAG ATC GGA AAA GAT ACA TTC ATT ACT CAC GAG CTG GCT GCC GAG GCA GCA CCA AAC CTG GTG GGG CCC CGT GAC TTT GTG AGC GTG CGC TGT GCC AAG CGC CGA GGC TCC ACC TGT GTG
Lys Ile Gly Lys Asp Thr Phe Ile Thr His Glu Leu Ala Ala Glu Ala Ala Gly Asn Leu Val Gly Pro Arg Asp Phe Val Ser Val Arg Cys Ala Lys Arg Arg Gly Ser Thr Cys Val>
```

FIG. 1(b)

```
                740                 760                 780                 800                 820                 840
CTG GCT GGC ATG GAC ACA GAC TTC GGG AAC ATG CCT GAG CAG AAG GGT GTC ATC AGG GCG GAG CAC GGT CCC ACT TGC ATG GTG CTT CAC CCG TTG GCT GGA AGT CCC TCT AAG ACC AAA
Leu Ala Gly Met Asp Thr Asp Phe Gly Asn Met Pro Glu Gln Lys Gly Val Ile Arg Ala Glu His Gly Pro Thr Cys Met Val Leu His Pro Leu Ala Gly Ser Pro Ser Lys Thr Lys>
                860                 880                 900                 920                 940                 960
CTT ACG TGG CTA CTC AGC ATC GAC CTC AAG GGG TGG CTG CCC AAG AGC ATC ATC AAC CAG GTC CTG TCC CAG ACC CAG GTG GAT TTT GCC AAC CAC CTG CGC AAG CGC CTG GAG TCC CAC
Leu Thr Trp Leu Leu Ser Ile Asp Leu Lys Gly Trp Leu Pro Lys Ser Ile Ile Asn Gln Val Leu Ser Gln Thr Gln Val Asp Phe Ala Asn His Leu Arg Lys Arg Leu Glu Ser His>
                980                1000                1020                1040                1060                1080
CCT GCC TCT GAA GCC AGG TGT TGA AGA CCA GCC TGC TGT TCC CAA CTG TGC CCA GCT GCA CTG GTA CAC ACG CTC ATC AGG AGA ATC CCT ACT GGA AGC CTG CAA GTC TAA GAT CTC CAT
Pro Ala Ser Glu Ala Arg Cys ...>
               1100                1120                1140                1160                1180                1200
CTG GTG ACA GTG GGA TGG GTG GGG TTC GTG TTT AGA GTA TGA CAC TAG GAT TCA GAT TGG TGA AGT TTT TAG TAC CAA GAA AAC AGG GAT GAG GCT CTT GGA TTA AAA GGT AAC TTC ATT

               1220                1240                1260                1280                1300                1320
CAC TGA TTA GCT ATG ACA TGA GGG TTC AGG CCC CTA AAA TAA TTG TAA AAC TTT TTT TCT GGG CCC TTA TGT ACC CAC CTA AAA CCA TCT TTA AAA TGC TAG TGG CTG ATA TGG GTG TGG

               1340                1360                1380                1400                1420                1440
GGG ATG CAT ACC ACA GGG CCT GAG AAG TCT TGG TTT ATG GGC TCA AGA ATG CCA TGC GCT GGC AGT ACA TGT GCA CAA AGC AGA ATC TCA GAG GGT CTC CTG CAG CCC TCT GCT CCT CCC

               1460                1480                1500                1520                1540                1560
GGC CGC TGC ACA GCA ACA CCA CAG AAC AAG CAG CAC CCC ACA GTG GGT GCC TTC CAG AAA TAT AGT CCA AGC TTT CTC TGT GGA AAA AGA CAA AAC TCA TTA GTA GAC ATG TTT CCC TAT

               1580                1600
TGC TTT CAT AGG CAC CAG TCA GAA TAA AGA ATC ATA ATT CAC ACC AAA AAA AAA AAA A
```

FIG. 2


NL Adrenal
NL Testis
Patient 1
Patient 2
NO RNA
Markers
kb
1.35
1.08
0.87
0.60


100 bp
Patient 1
R193 → Stop
5'
3'
Q258 → Stop
Patient 2
S1
AS 1

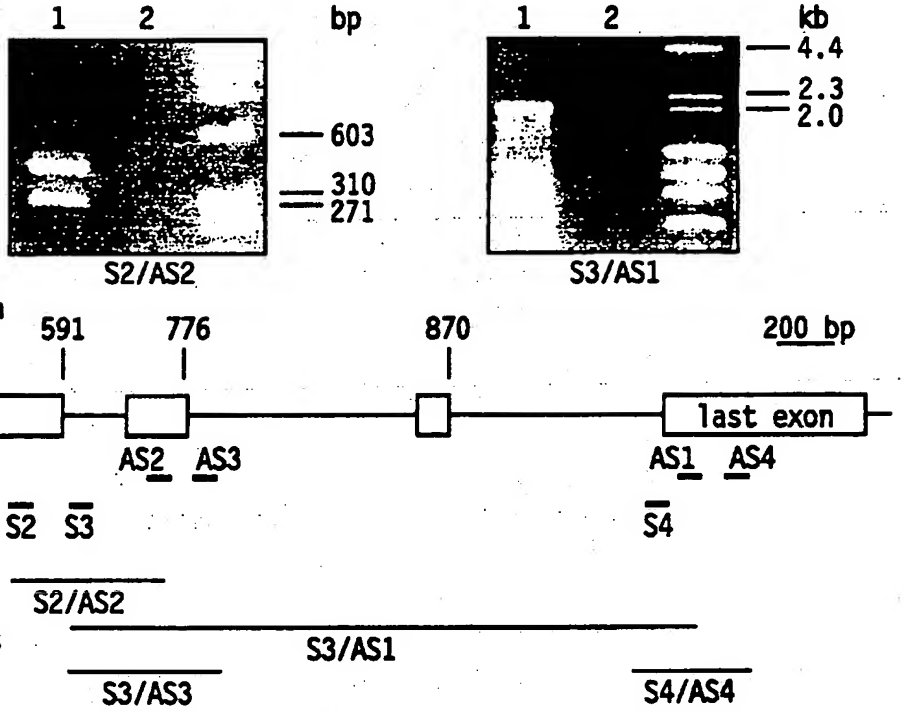
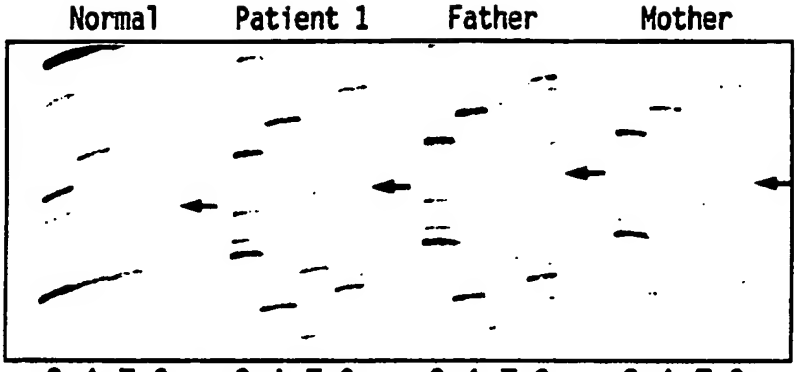

1 2
bp
603
310
271
S2/AS2
1 2
kb
4.4
2.3
2.0
S3/AS1
Position in cDNA
591
776
870
200 bp
last exon
PCR Primers
AS2 AS3
AS1 AS4
S2 S3
S4
PCR Products
S2/AS2
S3/AS1
S3/AS3
S4/AS4

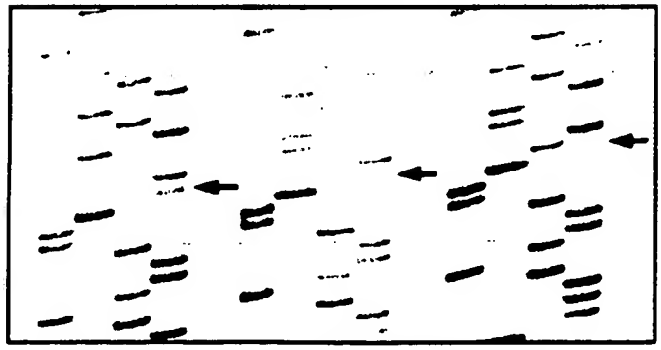
FIG. 3

FIG. 4A



G A T C	G A T C	G A T C	G A T C
Codon #	191	192 193	194 195
Amino Acid	Lys	Arg Arg	Gly Ser
Normal	5' AAG	CGC CGA	GGC TCC 3'
Patient 1		T	
Father		C/T	
Mother		C/T	

FIG. 4B



G A T C	G A T C	G A T C
Codon #	256 257	258 259
Amino Acid	Ile Asn	Gln Val
Normal	5'ATC AAC	CAG GTC 3'
Patient 2		T
Patient 3		T

FIG. 5A

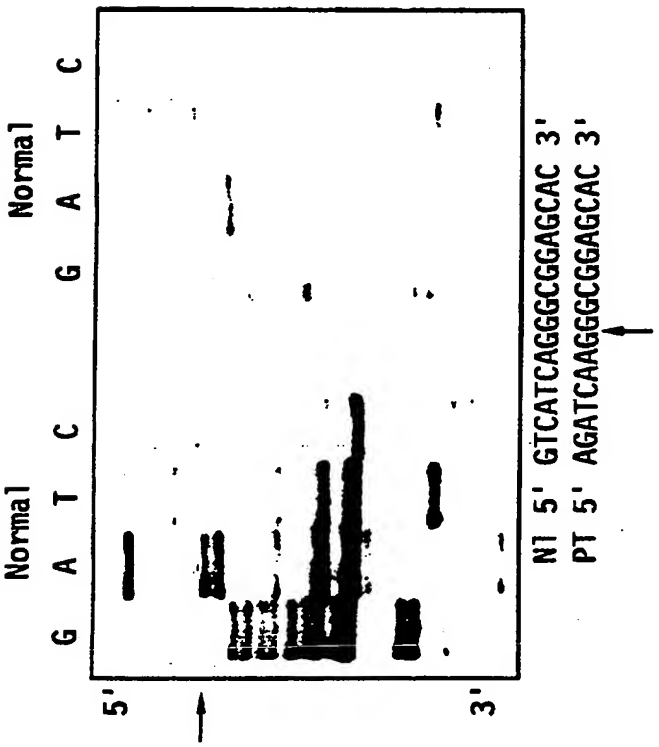


FIG. 5B

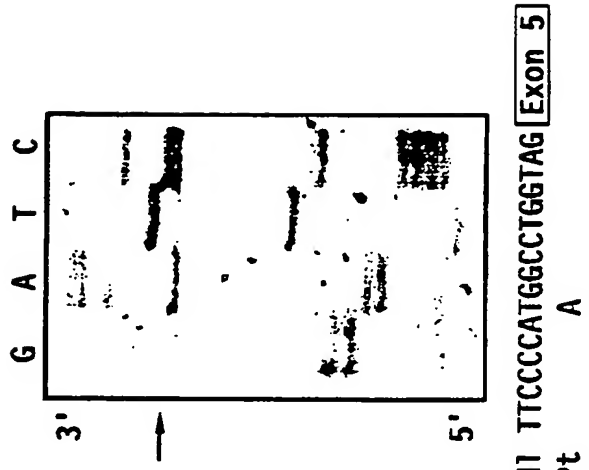


FIG. 5C

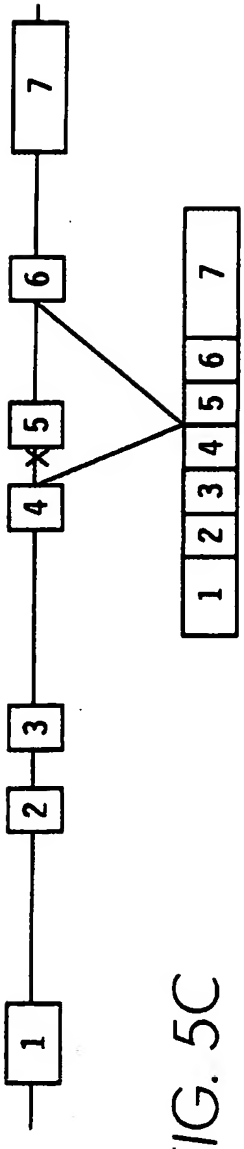


FIG. 6

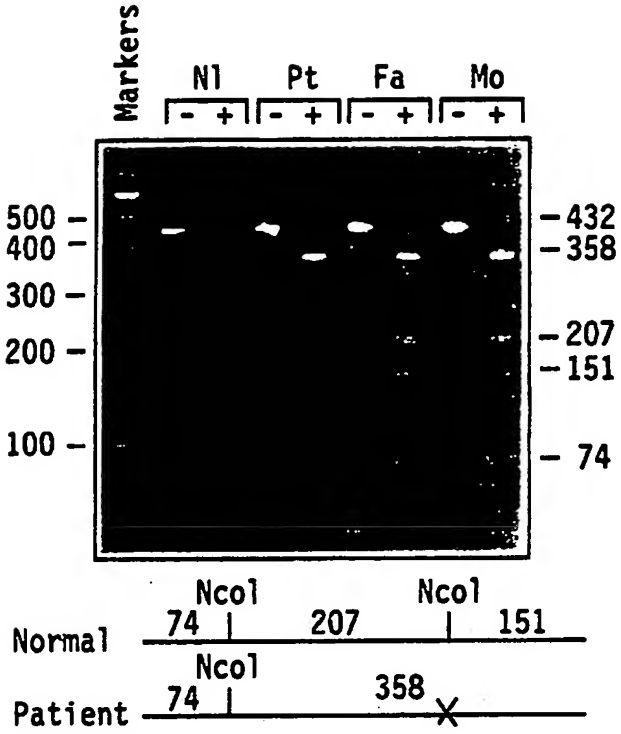


FIG. 8a

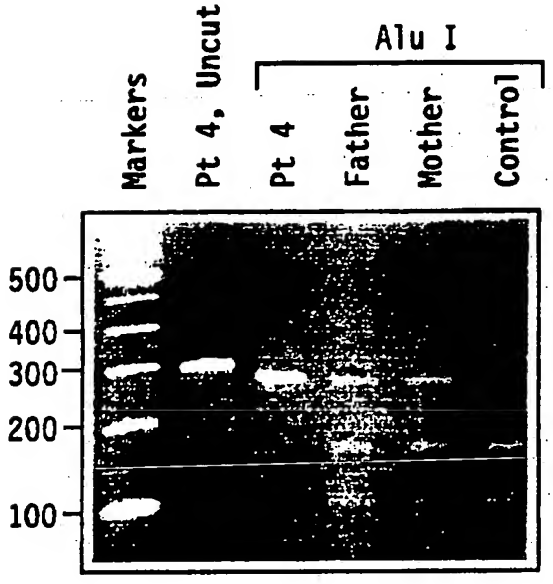


FIG. 8b

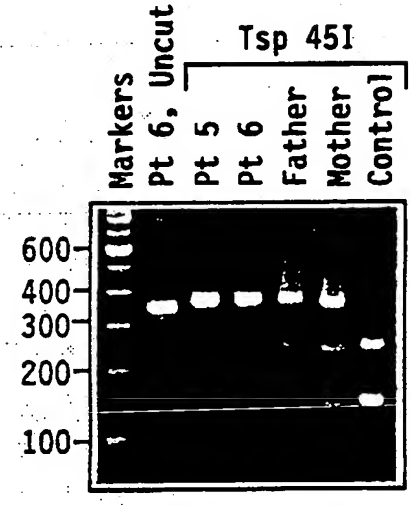


FIG. 7

AGCTTTCTGCACATACCAAGACCCCCAGCCCAGCTCACTCAGACAAAGCTACTGGCGGGAA -1239

AGTGTGAGGAAGGGTGTGGGCGTGGCCCAGGCCCTCCTCTTCTCTCTGCCGTATACTGATA -1178

Sp1

GGGCTGCCCCGCACCCCCCCCCGCCCCCCCGCGACTCAGCCACGAGAGGTCATCCTTGCTC -1117

CAGCACAAGACCCCTAAGAACCCTCACTCGGAACAGGACTTTGGGAAAGGTGGTTTTCTAT -1056

AAATAGATGAGTAAATAAATTGACAGTTGATATACCAAGCGTCCTGGGGCCGCAGGAGGAA -995

CTGTGTACAGATGGCTTGAAGGCCAGAGGCTTGGGTTGTTGTACTGGCCCTTCCACTGGCC -934

SF-1

AGCTGTTTGACCTTGAACAATCAAGTTCCACTCTGTGGACTTCAGGGTCCTCACCCAGAAG -855

AAGAGCAGCCATATGGTCTCTACTGCCTGGTAAACACCCTGGCTCACTCTCGCGAGATGGT -792

GGTTCTCCAAGTGTAGTGTGTAGTCCACACAACACCTGCATTGCAACCACTGGGTATTTTAT -731

TTATTTATTTAATTTATTTATTTAATTTATGATGGAGTCTCACTCTGTCGCCCAGGATGGA -670

GTGCAGTGGCACGATCTTGGCTTACTGCAACCTCTGCCTCCTGGGTTCAAGTGATTCTCAT -609

GCCTCAGCCTCCCGAGTAGCTGGGACTACAGGTGCCTGCCACATCACCCGGCTAATTTTTT -548

GTATTTTTAGTAGAGATGAGGTTTCACCATGTTGGCCATTCTGGTCTCGGACGCCTGACCT -487

CAAGTGATCTTCCCACCTCGGCCTCCCCAAGTTCTGGGGTTACAGGCGTAAACCACCGCCC -426

CTGGCCAAGGGGAGGTTTTTTCCTTTCCTTTTTTCCTTTTTCCTTTCCTTTCCCTCCTTTT -365

TTTTTTTTTTTTTTTTTTTTTTTTTAACACAGGTTTCTGAGCCTCAATTCCAGATCAGCTG -304

AGCCTGGAGTTTCTGAAGACAAGGGCTAGAAATCTGCACTTTAAAGTCTTGAAAACCACTG -243

TGTGCCTTCATCTAAGCTGCCCCTGCTTCTCTCCCCTCCATCCCTCGCCTGGCCCTGTCCT -182

Sp1

CCCTACTCTCCCCTGCACCCTCCCCCGCCCCAAGCTCCCCACAAACGGCCAAAGCAGCAGT -121

GTGAGGCAATCGCTCTATCCTTGACCCCTTCCTTTGCACAGTGAGTGATGGCGTTTTTATC -60

+1

TCCTGATGATGATGCACAGCCTTCAGCGGGGGACATTTAAGACGCAGAACACCAGGTCCA +1

GGCTGCAGCTGCGGGACTCAGAGGCGACTCAGAGGCGAAGCTTGAGGGGCTCAG +51

AAGGACGAAGAACCACCCTTGAGAGAAGAGGCAGCAGCAGCGGCGGCAGCAGCA +115

GCGGCAGCGACCCACCACTGCCACATTTGCCAGGAAACAATGCTGCTAGCGACA +169

M L L A T

FIG. 8c

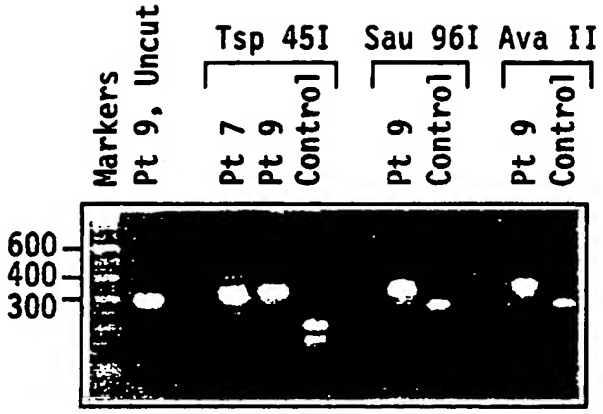

Markers
Pt 9, Uncut
Tsp 45I
Pt 7
Pt 9
Control
Sau 96I
Pt 9
Control
Ava II
Pt 9
Control
600
400
300

FIG. 8d

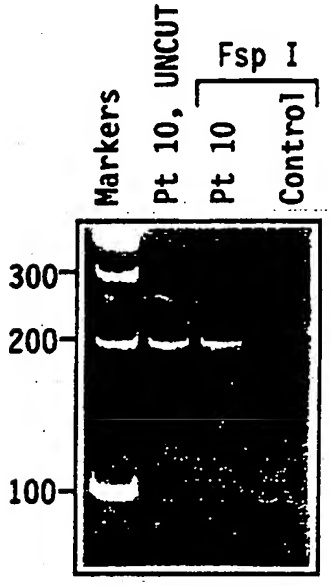

Markers
Pt 10, UNCUT
Fsp I
Pt 10
Control
300
200
100

FIG. 9A

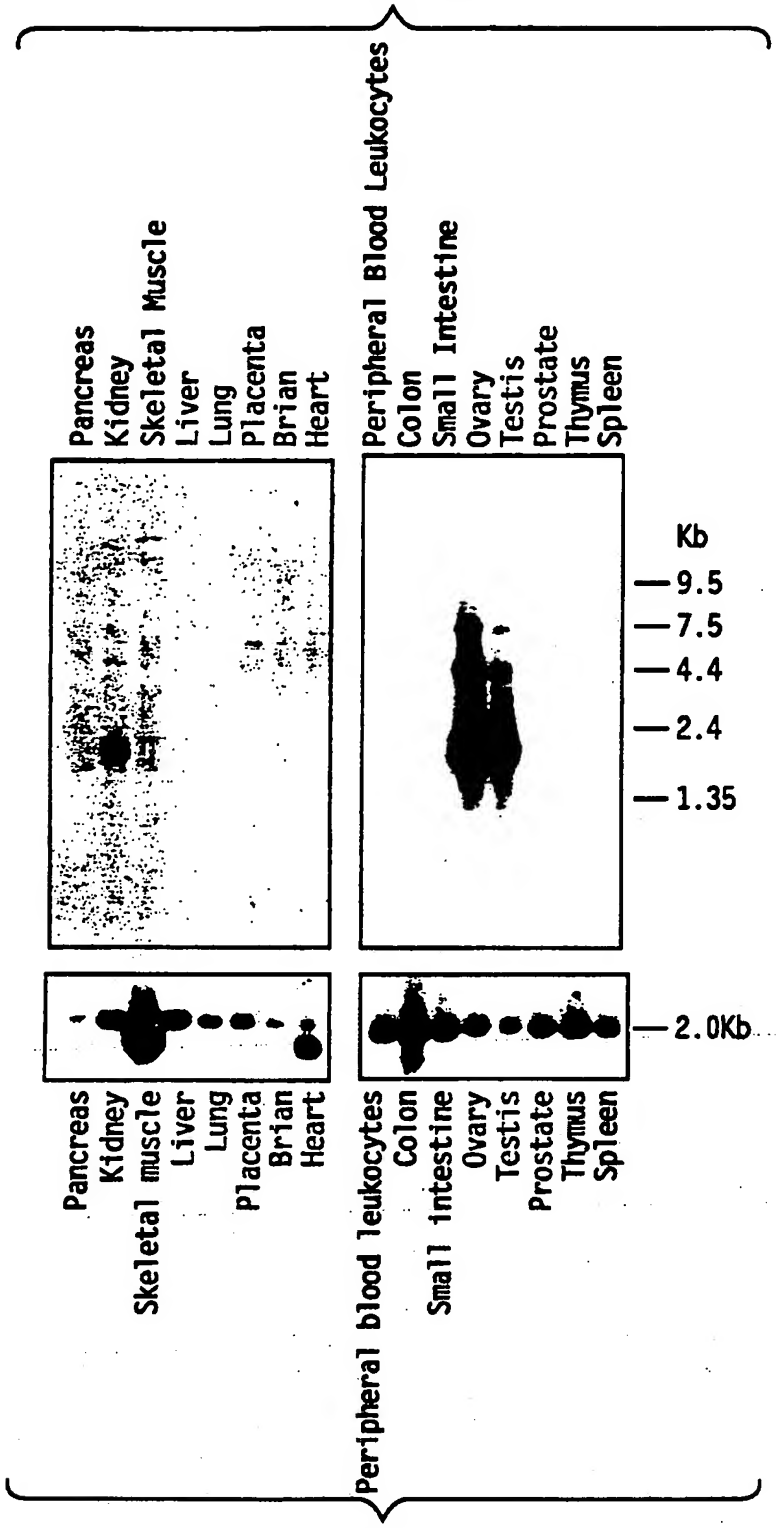

Pancreas
Kidney
Skeletal Muscle
Liver
Lung
Placenta
Brian
Heart
Peripheral Blood Leukocytes
Colon
Small Intestine
Ovary
Testis
Prostate
Thymus
Spleen
Kb
9.5
7.5
4.4
2.4
1.35
2.0Kb
Pancreas
Kidney
Skeletal muscle
Liver
Lung
Placenta
Brian
Heart
Peripheral blood leukocytes
Colon
Small intestine
Ovary
Testis
Prostate
Thymus
Spleen

FIG. 9B

FIG. 10


Trophoblast cells
Granulosa cells
Granulosa cells
- + - + - +
1.6 Kb
28 S rRNA

FIG. 11

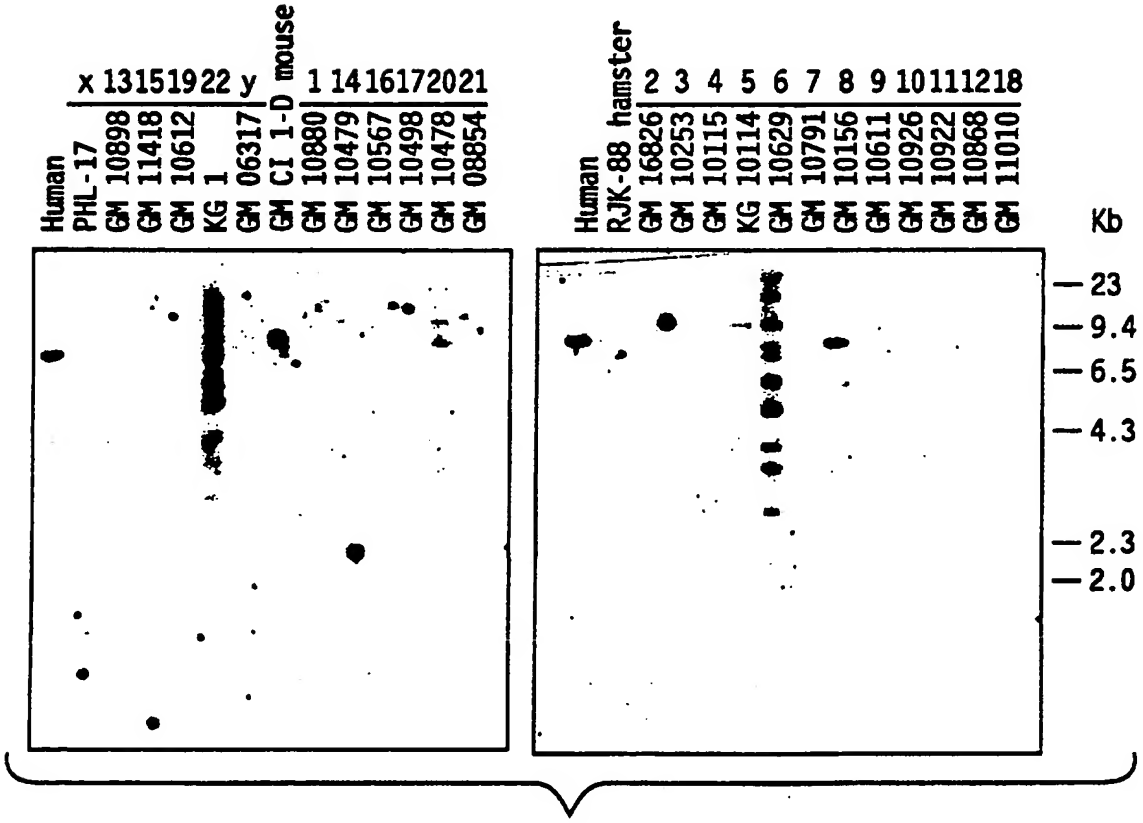

x 13151922 y 1 14161720 21
Human PHL-17
GM 10898
GM 11418
GM 10612
KG 1
GM 06317
GM CI 1-D mouse
GM 10880
GM 10479
GM 10567
GM 10498
GM 10478
GM 08854
Human RJK-88 hamster
GM 16826
GM 10253
GM 10115
KG 10114
GM 10629
GM 10791
GM 10156
GM 10611
GM 10926
GM 10922
GM 10868
GM 11010
2 3 4 5 6 7 8 9 10111218
Kb
23
9.4
6.5
4.3
2.3
2.0

FIG. 12A

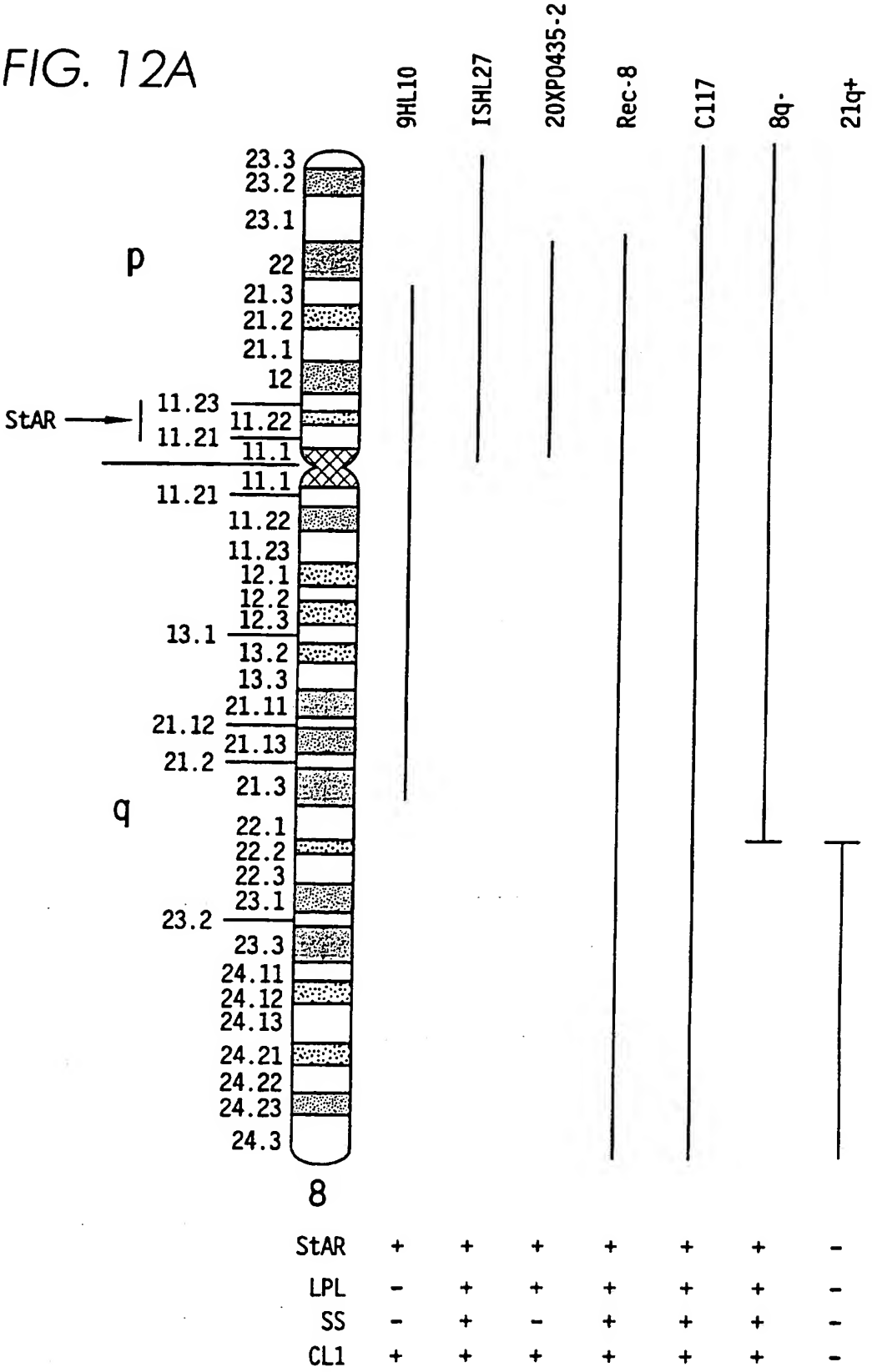

9HL10
ISHL27
20XP0435-2
Rec-8
C117
8q-
21q+
StAR
p
q
23.3
23.2
23.1
22
21.3
21.2
21.1
12
11.23
11.22
11.21
11.1
11.1
11.21
11.22
11.23
12.1
12.2
12.3
13.1
13.2
13.3
21.11
21.12
21.13
21.2
21.3
22.1
22.2
22.3
23.1
23.2
23.3
24.11
24.12
24.13
24.21
24.22
24.23
24.3
8
StAR + + + + + + -
LPL - + + + + + -
SS - + - + + + -
CL1 + + + + + + -

FIG. 13

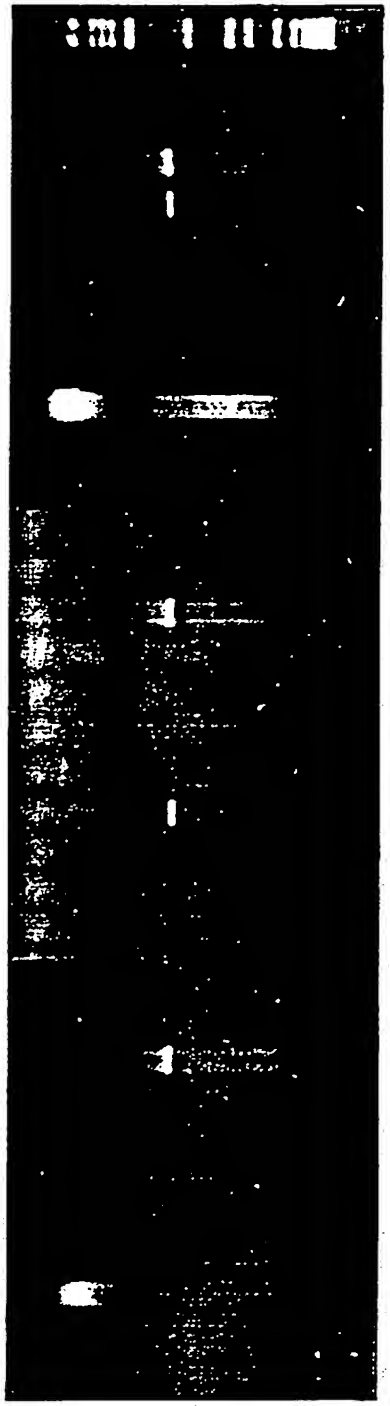


FIG. 12B

FIG. 14

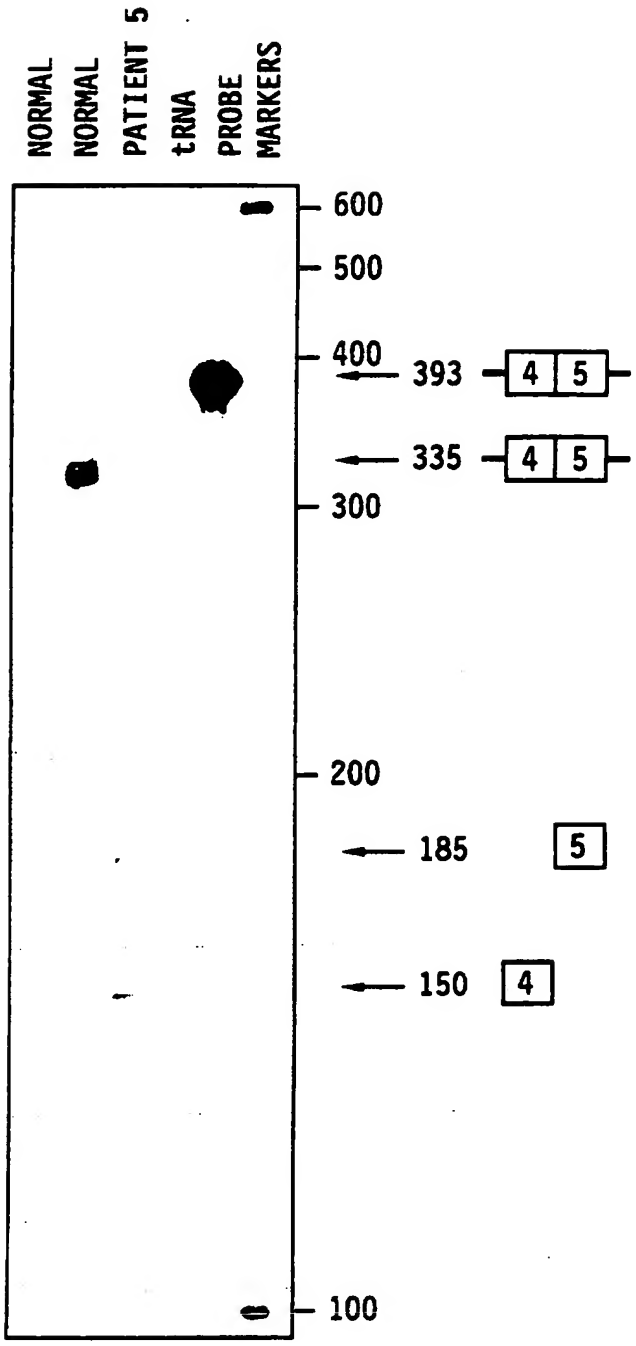

NORMAL
NORMAL
PATIENT 5
tRNA
PROBE
MARKERS
600
500
400
393
4
5
335
4
5
300
200
185
5
150
4
100

FIG. 15

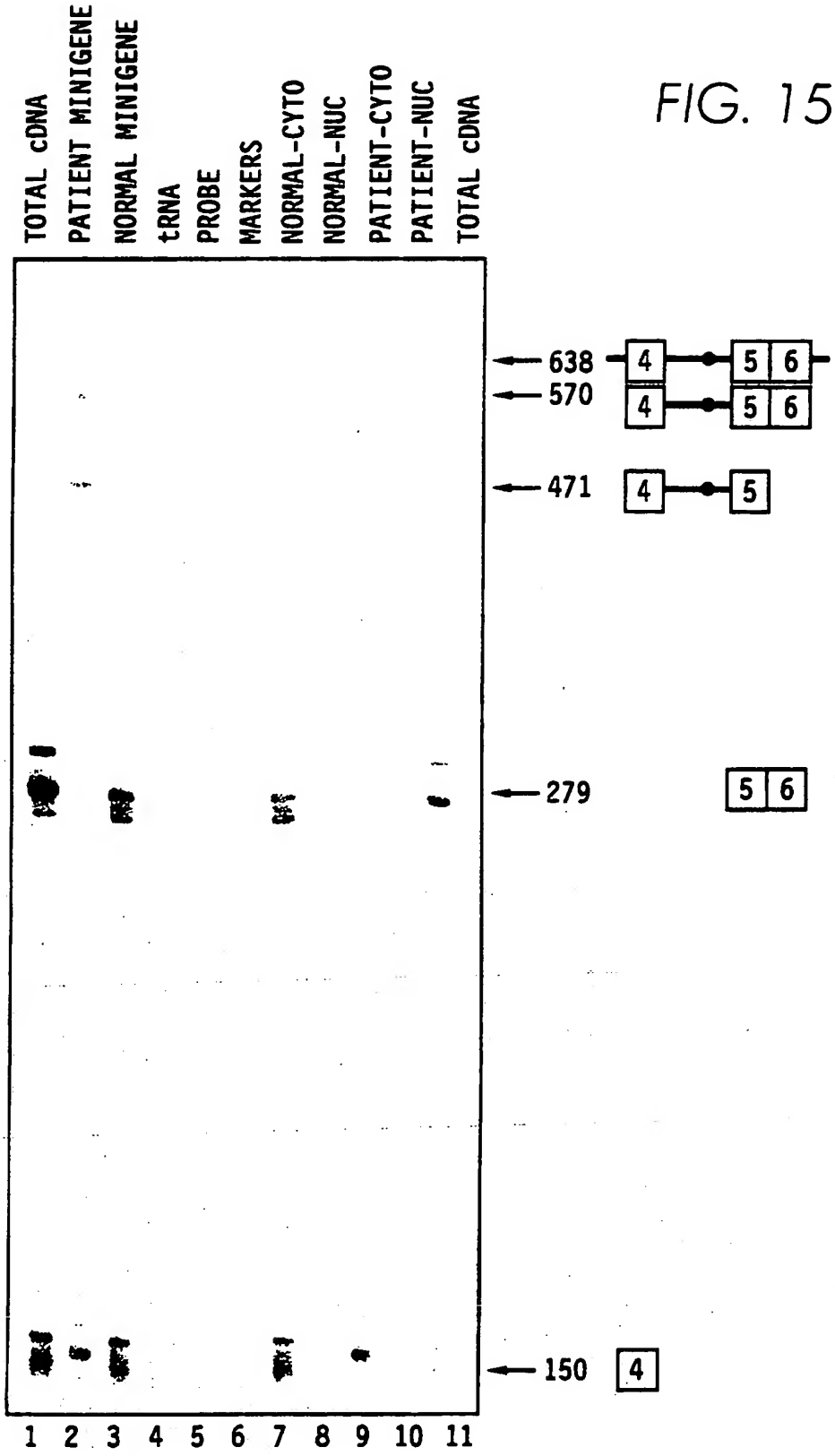
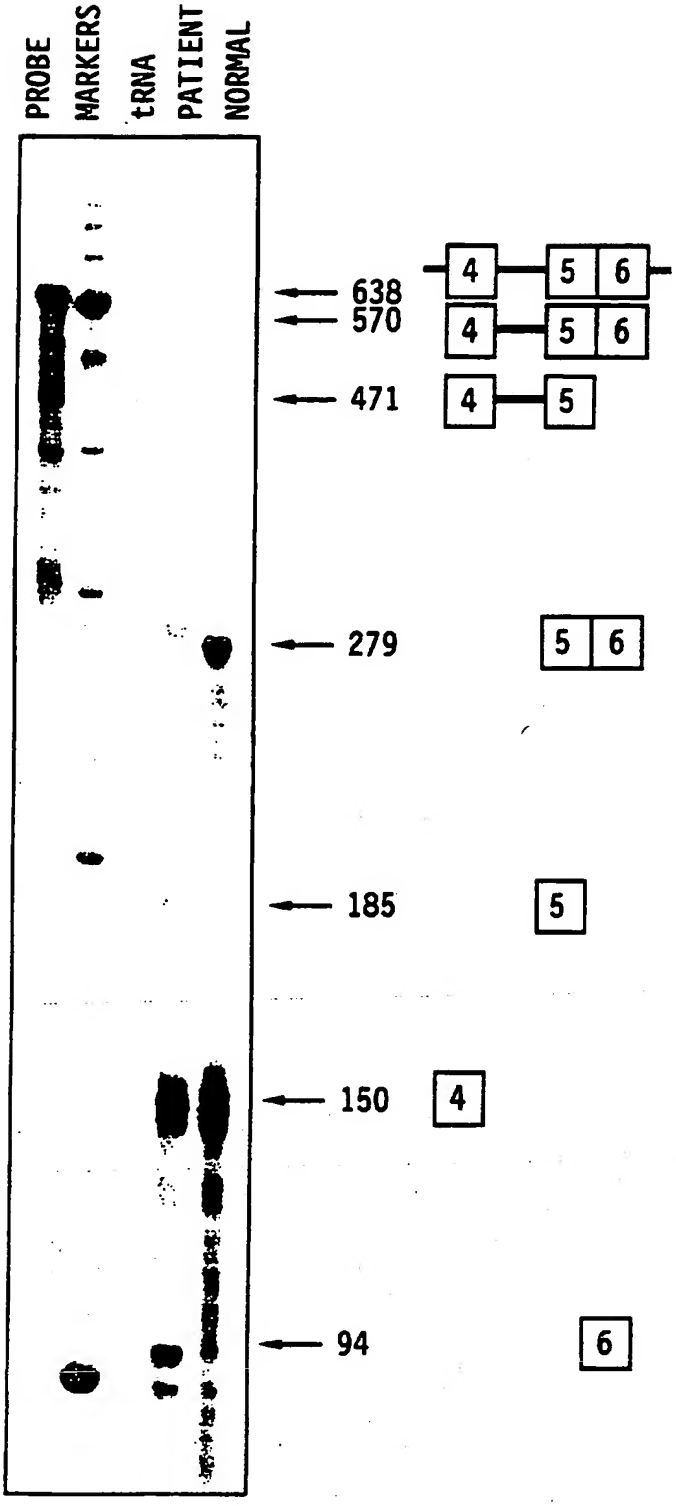

TOTAL cDNA
PATIENT MINIGENE
NORMAL MINIGENE
tRNA
PROBE
MARKERS
NORMAL-CYTO
NORMAL-NUC
PATIENT-CYTO
PATIENT-NUC
TOTAL cDNA
638
570
471
279
150
4
5
6
1 2 3 4 5 6 7 8 9 10 11

FIG. 16


PROBE
MARKERS
tRNA
PATIENT
NORMAL
638
570
471
279
185
150
94
4
5
6
4
5
6
4
5
5
6
5
4
6

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US96/03896

A. CLASSIFICATION OF SUBJECT MATTER
IPC(6) :C07H 21/02, 21/04; C12Q 1/68
US CL :536/23.1; 435/6
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 536/23.1; 435/6

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	STOCCO et al. The 30-kDa Mitochondrial Proteins Induced by Hormonal Stimulation in MA-10 Mouse Leydig Tumor Cells Are Processed from Larger Precursors. The Journal of Biological Chemistry. 15 October 1991, Vol. 266, No. 29, pages 19731-19738, see the entire document.	1-41
A	EPSTEIN et al. Regulation of Steroid Hormone Biosynthesis. The Journal of Biological Chemistry. 15 October 1991, Vol. 266, No. 29, pages 19739-19745, see the entire document.	1-41

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be part of particular relevance
"E" earlier document published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search
28 MAY 1996

Date of mailing of the international search report
02 AUG 1996

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231
Facsimile No. (703) 305-3230

Authorized officer
ETHAN WHISENANT
Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US96/03896

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	HUFFA et al. Congenital Adrenal Hyperplasia due to deficient cholesterol side-chain cleavage activity (20, 22-desmolase) in a patient treated for 18 years. Clinical Endocrinlogy. November 1985, Vol. 23, No. 5, pages 481-493, see the entire document.	1-41
Y	LIN et al. The Human Peripheral Benzodiazapine Receptor Gene: Cloning and Characterization of Alternative Splicing in Normal Tissues and in a Patient with Congenital Adrenal Hyperplasia. Genomics. December 1993, Vol. 18, pages 643-650, see especially pages 645-646, Figures 2-3.	1, 4-8
A	SAENGER et al. Prenatal Diagnosis of Congenital Lipoid Adrenal Hyperplasia. Journal of Clinical Endocrinology and Metabolism. January 1995, Vol. 80, No. 1, pages 200-205, see the entire document.	20-41